Structural and functional characterization of the RNA-binding proteins Loquacious and Brain tumor from *Drosophila melanogaster*

DISSERTATION ZUR ERLANGUNG DES DOKTORGRADES DER NATURWISSENSCHAFTEN (DR. RER. NAT.) DER FAKULTÄT FÜR BIOLOGIE UND VORKLINISCHE MEDIZIN DER UNIVERSITÄT REGENSBURG

Vorgelegt von

Leonhard Jakob

aus

Regensburg

Im Jahr 2016

Das Promotionsgesuch wurde eingereicht am:
18.11.2016

Die Arbeit wurde angeleitet von:
Prof. Dr. Gunter Meister

_____

(Leonhard Jakob)

Zwei Dinge sind zu unserer Arbeit nötig:
Unermüdliche Ausdauer und die Bereitschaft, etwas, in das man viel Zeit und Arbeit gesteckt hat, wieder wegzuwerfen.

Albert Einstein

Parts of this thesis have been published in the following articles:

Jakob, L., Treiber, T., Treiber, N., Gust, A., Kramm, K., Hansen, K., Stotz, M., Wankerl, L., Herzog, F., Hannus, S., Grohmann, D., Meister, G., 2016. Structural and functional insights into the fly microRNA biogenesis factor Loquacious. RNA 22, 383–96.

Loedige, I., Jakob, L., Treiber, T., Ray, D., Stotz, M., Treiber, N., Hennig, J., Cook, K.B., Morris, Q., Hughes, T.R., Engelmann, J.C., Krahn, M.P., Meister, G., 2015. The Crystal Structure of the NHL Domain in Complex with RNA Reveals the Molecular Basis of Drosophila Brain-Tumor-Mediated Gene Regulation. Cell Rep. 13, 1206–20.

Furthermore, I contributed to the following articles:

Hasler, D., Lehmann, G., Murakawa, Y., Klironomos, F., Jakob, L., Grässer, F.A., Rajewsky, N., Landthaler, M. and Meister, G., 2016. The Lupus Autoantigen La Prevents Mis-channeling of tRNA Fragments into the Human MicroRNA Pathway. Mol. Cell 1–15.

Hauptmann, J., Schraivogel, D., Bruckmann, A., Manickavel, S., Jakob, L., Eichner, N., Pfaff, J., Urban, M., Sprunck, S., Hafner, M., Tuschl, T., Deutzmann, R., Meister, G., 2015. Biochemical isolation of Argonaute protein complexes by Ago-APP. Proc. Natl. Acad. Sci. U. S. A. 112, 11841–5

Parts of this thesis have been presented on the following conference:

The 18th annual meeting of the RNA Society. Davos, Switzerland

Title of poster: Establishment of a Fluorescence Cross Correlation (FCCS) based assay to measure Argonaute - siRNA interactions

Coordinates and structural-factors have been uploaded to the RCSB protein-databank (www.pdb.org) with the following submission codes:

4X8W: dsRBD3 of Loquacious

4ZLR: Structure of the Brat-NHL domain bound to consensus RNA motif

Abbreviations

A	
A	Ampere or absorption
aa(s)	Amino acid(s)
Ab	Antibody
Ala (A)	Alanine
Alexa Fluorophore 647	AF647
Amp	Ampicillin
Asn (N)	Asparagine
Asp (D)	Aspartate
Arabidopsis	*Arabidopsis thaliana*
Arg (R)	Arginine
AU	Absorption unit
B	
Bp(s)	Base pair(s)
C	
Cam	Chloramphenicol
CDS	Coding sequence
Cys (C)	Cysteine
Cv	Column volume
D	
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
ds	Double stranded
dm or *Drosophila*	*Drosophila Melanogaster*
dNTP	Deoxynucleoside triphosphate
E	
E. coli	*Escherichia coli*
EDC	1-ethyl-3-(3-dimethyl-aminopropyl)-carbodiimid
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electrophoretic mobility shift assay
EtBr	Ethidium bromide
F	
FBS	Fetal bovine serum
FCCS	Fluorescence cross-correlation spectroscopy
FH	FLAG/HA
G	
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GFP	Green fluorescent protein
Gln (Q)	Glutamine
Glu (E)	Glutamate
Gly (G)	Glycine
GST	Glutahione-S-transferase
H	
Hepes	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HEK	human embryonic kidney
His	Histidine (H)
I	
Ile (I)	Isoleucine
IMAC	Immobilized metal ion affinity chromatography
IP	Immunoprecipitation
IPTG	Isopropyl-β-D-thiogalactoside

K	
Kan	Kanamycin
L	
LB	Lysogeny broth
Leu (L)	Leucine
Lys (K)	Lysine
M	
m	Meter or milli
M	Molar
m^7G	7-methylguanosine
Met (M)	Methionine
min	Minute
miRNA	Micro RNA
mRNA	Messenger RNA
MS	Mass spectrometry/spectrometric-
MW	Molecular weight
N	
nc	Non-coding
NCS	Non-crystallographic symmetry restraints
nt(s)	Nucleotide(s)
O	
o/n	Over night
P	
PABP	Poly(A) binding protein
PCR	Polymerase chain reaction
PEG	Poly ethylene glycol
Phe (F)	Phenylalanine
piRNA	Piwi-interacting RNA
Pre-miRNA	Precursor miRNA
Pri-miRNA	Primary miRNA
Pro (P)	Proline
PTGR	Post-transcriptional gene regulation
Q	
qRT-PCR	Quantitative real time PCR
R	
RBD	RNA binding domain
RISC	RNA induced silencing complex
RLC	RISC loading complex
RNA	Ribonucleic acid
RRM	RNA-recognition motif
RT	Room temperature
S	
S. cerevisiae	*Saccharomyces cerevisiae*
SDS	Sodium dodecylsulphate
SEB	Size exclusion buffer
SEC	Size exclusion chromatography
SeMet	Selenomethionine
Ser (S)	Serine
snRNA	small nuclear RNA
ss	Single-stranded
T	
Thr (T)	Threonine
TRBP	Trans-activating response RNA-binding protein
Tris	Tris(hydroxymethyl)-aminomethane
Trp (W)	Tryptophan

Trx	Thioredoxin
Tyr (Y)	Tyrosine
U	
U	Uracil or Unit
UTR	Untranslated region
V	
V	Volt
Val (V)	Valine
W	
W	Watt
wt	Wildtype
X	
X. laevis	*Xenopus laevis*

Contents

Abstract

Post-transcriptional gene regulation (PTGR) processes play a major role in eukaryotic and prokaryotic gene expression and mainly act on the mRNA level. Therefore, among other cellular processes, PTGR allows cells to react on constantly changing environmental influences, by the ability to rapidly adjust internal protein levels. As key players in such regulatory processes, RNA-binding proteins (RBPs) interact with mRNAs to regulate their maturation, localization, translation and decay. Thereby RBPs can directly recognize their RNA target sequences via specific contacts with the RNA bases, for example. In other cases, the RBP binds to a short fragment of nucleic acids, which guides the RBP-RNA complex to a complementary target sequence via Watson-Crick base pairing with the target RNA. Examples for such RNA guides are microRNAs (miRNAs), short-interfering RNAs (siRNAs) and P-element-induced wimpy testis (Piwi)-interacting RNAs (piRNAs).

MiRNAs represent a class of small non-coding RNAs (ncRNAs), which are highly conserved in biogenesis and function among eukaryotes. After processing of the miRNA from double stranded precursors, by a variety of different proteins, including the RNase III enzyme Dicer and its double-stranded RNA binding protein (dsRBP) partners, mature miRNAs assemble with Argonaute (Ago) proteins into so called miRNA-induced silencing complexes (miRISCs). In this study, the processing of the *Drosophila melanogaster* (*Drosophila*) Dicer-1 (dmDcr-1) in complex with its dsRBP partner Loquacious (Loqs) has been structurally and functionally investigated. Loqs contains three different dsRBDs and the crystal structure of the third dsRNA binding domain (dsRBD) reveals specific structural elements that are responsible for the interaction with dmDcr-1 and a linker preceding the dsRBD3 strongly enhances the affinity of the binding. Furthermore, the Loqs-dmDcr-1 complex revealed a 1 to 1 stoichiometry identified by single molecule spectroscopy but formed dimers in the absence of dmDcr-1. In consideration of the conservation of critical sequence elements, it is likely that dimerization might comprise a general feature of dsRBPs in gene silencing.

Ago proteins, as protein-binding partners of miRNAs, are key players in miRNA-mediated gene-silencing processes. For their correct function, they require a strict regulation that is supposed to be mainly achieved by phosphorylation of specific residues within the sequence of Ago proteins. In mass spectrometry (MS) screens, several potential regulatory sites have been reported in literature, for example Y593 and S798 within the sequence of Ago2. To be able to investigate the kinetical details behind Ago-loading processes of wildtype (wt) and mutant Ago proteins, a confocal fluorescence cross-correlation spectroscopy (FCCS) assay was established. Thereby, the binding of a fluorescently labeled siRNA duplex to GFP-tagged Ago1-4 and respective mutants was investigated. Furthermore kinetic data from different mutations of the potential post-translational regulatory phospho-sites was measured, providing information on the binding of a

siRNA 5’ end within the 5’ binding pocket. Furthermore, the established FCCS assay provides a valuable tool for the future analysis of any mutation within the sequence of an Ago protein and its influence on Ago-loading processes.

In the second part of this thesis, the *Drosophila* RBP brain tumor (Brat), which belongs to the conserved NHL-domain containing tripartite motif (TRIM) protein family, has been structurally characterized. Brat is responsible for the differentiation of neuronal stem cells by repressing the expression of self-renewal factors. Thereby, Brat directly contacts a specific target sequence motif on mRNAs, coding for self-renewal factors and is supposed to function as a translational repressor. The molecular details for the interaction of Brat with its RNA targets, however, were unknown. Therefore, the crystal structure of the Brat NHL-domain in complex with its consensus target sequence has been solved at 2.3 Å resolution. The high resolution protein-RNA structure reveals several features of sequence-specific substrate contacts of the Brat-NHL domain to its consensus RNA motif, adding the β-propeller fold as a novel member to the growing family of RBDs.

I. Introduction

For any living cell from all three kingdoms of life, the genetic code is stored within the sequence of its own deoxyribonucleic acid (DNA). Since the sequencing of the human genome (Lander et al., 2001) it is known that sequences for protein-coding genes make up only 25 % (Introns and Exons) of the total DNA. The proportion of Exons, which actually encode for polypeptides, is even less with only ~1 % (Figure 1.1 A). The remaining part belongs to other intergenic (22 %) or repetitive DNA (53 %).

To make the information accessible for the cell, it has to be transcribed into a ribonucleic acid (RNA) macromolecule by a DNA-dependent RNA polymerase. In eukaryotes, three different polymerases have been identified (Roeder and Rutter, 1969) and (Roeder and Rutter, 1970), which are distributed to different loci within the nucleus and are responsible for the synthesis of diverse RNA species. RNA polymerase I (PolI) is located in the nucleolus where its main function is to transcribe the 28S, 18S and 5.8S precursor ribosomal RNAs (pre-rRNAs), which are then further processed to mature rRNAs. Although only 6 species exist among that class of RNA, rRNAs represent the highest abundance referred to the total RNA mass in the cell (Figure 1.1 B).

RNA polymerase III (PolIII) is mainly required for the synthesis of transfer RNAs (tRNAs), which represent the second most abundant class of RNAs (Figure 1.1, B), the 5S rRNA and several other shorter RNA species (e.g U6-, RNaseP- and MRP-, 7SL-, 7SK-, Vault- Y-RNAs, reviewed in Dieci et al. (2007)). The main function for RNA polymerase II (PolII) is the transcription of protein-coding genes into a precursor messenger RNA (pre-mRNA) transcript. Additionally, PolII is required for the transcription of noncoding RNAs, such as small nuclear RNAs (snRNAs), miRNAs and piRNAs.

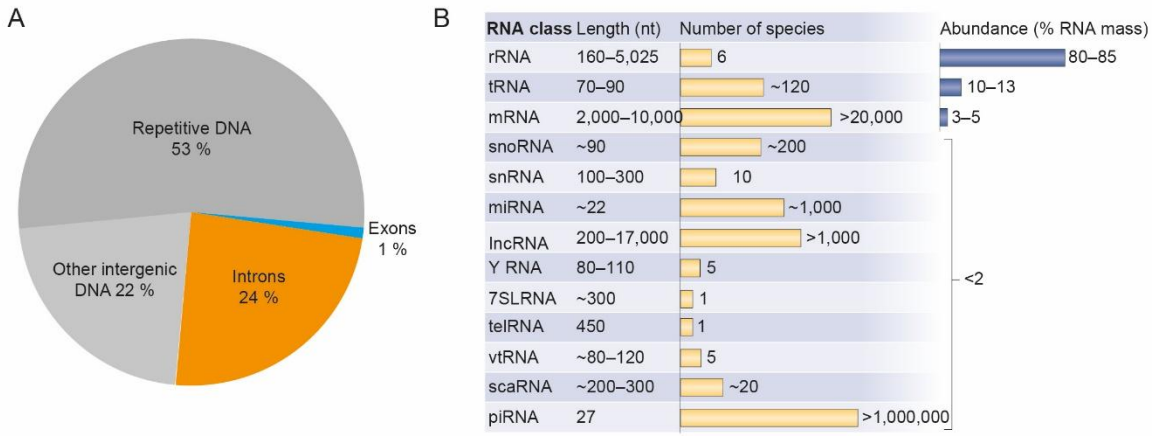


Figure 1.1: Schematic overview of the genomic DNA and RNA classes in humans. (**A**) Genes consisting of Introns (24 %) and Exons (1 %) only constitute a small fraction of the human genome. The main part of genomic DNA consists of repetitive (53%) and other intergenic DNA (22%). (**B**) Overview of different classes of the RNA transcriptome. Adapted from Krebs et al., (2014) and Jankowsky and Harris, (2015).

Soon after synthesis, the pre-mRNAs are further processed by a specific splicing machinery, which removes intron sequences and retains exon sequences containing the genetic information. This process is not linear, resulting in only one particular splice isoform of a protein, but can exhibit a variety of different isoforms by alternative splicing pathways. The protein coding regions located on the mRNA are flanked by two untranslated regions (UTR). The 5' UTR ranges from the start point of transcription to the start codon of translation and harbors the ribosomal binding site. Additionally, 5' UTRs are binding platforms for various regulatory proteins, which can influence the stability of the mRNA or regulate translational activity. The 5' end is capped with 7-methylguanosine (m^7G), which protects the mRNA from 5'->3' exonucleolytic degradation. The 3' UTR starts directly after the stop codon and harbors a polyadenylation signal, which then leads to polyadenylation of the 3' end of the mRNA. The poly(A) tail is bound by the poly(A) binding protein (PABP) that protects the end from degradation and has various additional regulatory functions. Compared to the 5' UTR, the 3' UTR is more frequently bound by a multitude of regulatory proteins which either bind directly or indirectly to that region. As protein synthesis is one of the most important biochemical processes this step needs to be strictly regulated to provide the possibility to react on various environmental influences.

1. PTGR mediated by RBPs

PTGR can occur on many RNA transcripts in living organisms. Therefore, cells are provided with an extensive repertoire of specific RBPs, which are indispensable for the biogenesis, stability, function, transport and cellular localization of functional RNA precursors. In prokaryotes, where transcription is directly coupled to translation, the amount of RBPs is rather limited. In eukaryotes, however, transcription and translation are spatially separated (nucleus and cytoplasm), which creates an additional layer of gene regulation. Hence, the number of RBPs increases strongly with the complexity of the organism, which might derive from an evolution of highly specific processes to fine-tune gene expression (Anantharaman et al., 2002). A census of human RBPs ends up with 1542 different proteins that interact with all known classes of RNAs (Gerstberger et al., 2014). PTGR is not restricted to mRNA-associated processes in general, but can also be found in processes acting on non-coding RNAs (ncRNAs). In the Online Mendelian Inheritance in Man (OMIM) database, which contains information on human genes and genetic disorders, there are ~150 annotated RBPs linked to human diseases, where only one third is associated with mRNA binding, whereas two third of RBPs act on ncRNAs (Gerstberger et al., 2014; Hamosh et al., 2005).

Gene expression of protein coding genes involves transcription, translation and turnover of mRNAs and proteins. The expression of a certain mRNA correlates largely with the abundance of the encoded protein. However, there is a strong deviation between half-lives of mRNAs and

proteins (Schwanhäusser et al., 2011). This emphasizes the meaning of RBP-mediated post-transcriptional mechanisms that are required for the regulation of mRNAs. Although many RBPs have been identified, only a limited number of RNA binding domains have been structurally characterized and little is known how binding selectivity and specificity are accomplished.

Although RNA molecules are single stranded, they are able to form intra- and intermolecular double stranded structures by specific base pairing. For PTGR these structures need to be recognized by RBPs, which therefore comprise a repertoire of several different RBDs, to specifically identify their target sequences on RNAs. The most prominent representatives for either single stranded (ss) or double stranded (ds) RNA recognition will be introduced within the following paragraphs.

1.1 RBP-RNA interactions

RBPs are highly specialized proteins for the recognition of certain features from their target RNA macromolecules. The driving forces behind this interaction can be as diverse as electrostatic interactions of positively charged protein side chains with the negative phosphate backbone of the RNA, stacking interactions of aromatic side chains with the aromatic purine or pyrimidine bases and specific hydrogen bonds of donor and acceptor groups between the protein and the RNA.

1.1.1 Electrostatic interactions

Electrostatic interactions play an important role for the interaction between RBPs and their target RNAs. Nucleic acids are highly negatively charged molecules, which can be coordinated by positively charged protein-surfaces. Salt bridges between two contrary charged molecules strongly contribute to the binding affinity and are present in almost every RBP to tightly anchor its respective RNA substrate. In some cases, the arrangement of positively charged amino acids allows the RBP to recognize characteristic shapes of tertiary RNA structures (Oberstrass et al., 2006; Stefl et al., 2005). The contribution of electrostatic interactions for sequence-specific recognition of RNA molecules, however, is rather limited, as the distribution of charges from ssRNA is made up from its sugar-phosphate backbone and therefore independent from its nucleotide sequence. Therefore, electrostatic interactions represent the major part of protein-RNA contacts in proteins, which function sequence independently. Some examples are RNA polymerases (Westover et al., 2004), the DEAD-box protein Vasa (Sengoku et al., 2006) or certain viral nucleoproteins (Albertini et al., 2006).

1.1.2 π stacking interactions

Another possibility for the interaction of RBPs with ssRNA is by quadrupole interactions of two delocalized electron systems with the positively charged carbon backbone, a so called π-π interaction. Thereby aromatic side chains of phenylalanine, tyrosine, tryptophan and histidine in RBDs of different RBPs are able to form π-stacking interactions with bases of nucleic acids in an parallel orientation with an average distance of ~3.3 Å in between the planes (Šponer and Hobza, 2003). For aromatic heterocycles it has been shown that purines are better stacking partners then pyrimidines and that stacking of aromatic systems is also strongly influenced on the substitution of the aromatic system (Sigrid D. Auweter et al., 2006; Cozzi and Siegel, 1995). Statistical comparisons of protein-RNA complexes, however, indicate that stacking interactions are more or less equally distributed among all four bases and are mediated most often by phenylalanines on the protein side (Allers and Shamoo, 2001). This indicates that aromatic π-stacking interactions alone rather contribute to the binding affinity then to sequence-specificity. Single point-mutations from aromatic amino acids, which are involved in π-stacking interactions, can influence the binding affinities by the magnitude of ~0.2-3x10^3 (Sigrid D Auweter et al., 2006; Deardorff and Sachs, 1997; LeCuyer et al., 1996).

Another way for the coordination of aromatic heterocycles of the nucleobases, which is similar to π-stacking interactions, is the formation of cation-π interactions that are mediated by arginine (Arg) residues. Thereby, the planar, positively charged guanidinium group of Arg interacts with the aromatic system of RNA bases with the preferential order of U,A,C > G, which allows a binding in a more sequence specific manner (Allers and Shamoo, 2001; Gallivan and Dougherty, 1999; Hudson et al., 2004).

1.1.3 Hydrogen bonds

Hydrogen bonds play the most important part in specific RNA-sequence recognition of RBPs. In a hydrogen bond, two electronegative atoms share a common proton, which is provided by a donor group and transferred to an acceptor group with a free electron pair that is capable of accommodating a proton. From the protein part, the side chains of tryptophan (Trp), lysine (Lys) and arginine (Arg) act as hydrogen-donors, whereas the side chains of aspartate (Asp), and glutamate (Glu) as hydrogen-acceptors and side chains of tyrosine (Tyr), serine (Ser), cysteine (Cys) threonine (Thr), asparagine (Asn), glutamine (Gln) and histidine (His) can act as both, hydrogen-donors and -acceptors. Additionally to side chains, the main chain of proteins can also contribute to the formation of hydrogen bonds with nucleotides by its peptide bond, consisting of a hydrogen-donor (N-H) and -acceptor group (C=O). The composition of hydrogen-donor and acceptor groups varies within different positions of the bases and thus allows a specific interaction with the protein backbone by a respective network of hydrogen bonds. The contribution of each

single hydrogen bond to the specificity is rather weak. However, taken all hydrogen-bonds together gives a substantial proportion to binding affinity between different RNA sequences and a protein and thus constitutes its sequence specificity (Sigrid D. Auweter et al., 2006). Different factors can influence the strength of a hydrogen bond. Hydrogen bonds, which are embedded in an hydrophobic surrounding, are reinforced compared to those which are located in more accessible parts of the protein (Deechongkit et al., 2004). Additionally, the exact relative geometrical orientation of the hydrogen donor in relation to the acceptor has a strong impact on the binding energy of a single hydrogen bond (Chen et al., 2004).

1.2 Some examples for ssRBDs

RBPs are commonly classified according to their content of RBDs, which are responsible for their binding preferences to a target RNA. RBDs are highly conserved across bacteria, archaea and eukaryotes (Gerstberger et al., 2014). From the total number of known RBDs, most members are associated in mRNPs, which reflects the rapid expansion of mRNA-related processes in the evolution of higher eukaryotes, like alternative splicing and polyadenylation (reviewed in: Chen and Manley, 2009; Keren et al., 2010).

The most abundant and best characterized RBD is the RNA-recognition motif (RRM) (Figure 1.2 A). It is composed of 80-90 amino acids, which form a four-stranded anti-parallel β-sheet flanked by two α-helices, resulting in a βαββαβ topology (Oubridge et al., 1994). RNA contacts are mediated at the two central strands of the β-sheet by two conserved sequence stretches, namely RNP1 with the sequence consensus K/R-G-F/Y-G/A-F/Y-V/I/L-X-F/Y and RNP2 with the sequence consensus V/I/L-F/Y-V/I/L-X-N/L. Thereby RRMs contact the phosphodiester backbone of RNAs by salt bridges with the conserved Arg or Lys residue and form specific interactions with the aromatic amino acids by stacking interactions. The total number of nucleotides which were bound by a single RRM ranges from a minimum of two in the cases of e.g. CBP20 (Calero et al., 2002; Mazza et al., 2002) to a maximum of eight for U2B0 (Price et al., 1998). Others than the name might let expect, RRMs do not only interact with RNAs, but can also mediate protein-protein contacts, depending on the positions of the α-helices within the topology (Crichlow et al., 2008; Tripsianes et al., 2014).

A second domain, which is found regularly in RBPs, is the K-homology (KH) domain. This unique RNA binding fold was first described in the heterogeneous nuclear ribonucleoprotein (hnRNP) K and found to be conserved among eukaryotes, eubacteria and archaea (Siomi et al., 1993). The domain consists of around 70 aas with a functionally important signature sequence of (I/L/V)IGXXGXX(I/L/V) close to the center of the domain and is found to bind both RNA and DNA target sequences (Braddock et al., 2002; Lewis et al., 2000). The KH domains consist of alternating α-helices and β-strands and according to their topology can be assigned to two

subfamilies. Type I (Figure 1.2 B) has a βααββα topology and type II (Figure 1.2 C) a αββααβ topology (Grishin, 2001).

Zinc finger proteins (Figure 1.2 D), which are also strongly involved in the interaction with DNA, are found to interact with both, ss- and dsRNA as well. The classification of zinc fingers is accomplished according to the coordination of a zinc ion by Cys and His residues within the RBD (CCCH, CCHC and CCHH).

The Piwi-Ago-Zwille (PAZ) and Piwi domains (Figure 1.2 E and F) are two kind of RBDs, which are generally found in the miRNA biogenesis pathway and are able to specifically recognize their substrates in a rather sequence independent manner. The PAZ domain consists of 110 amino acids and contains a β-barrel that is structurally similar to the oligonucleotide/oligosaccharide binding (OB) domain, where the nucleic acid binding site is located in a hydrophobic cleft between the β-barrel and a conserved module comprising strands β3, β4 and helix α3 (Lingel et al., 2003; Schirle and MacRae, 2012). In the RNase III-type enzyme Dicer, the PAZ domain is responsible for substrate recognition and positioning of the RNA prior to cleavage (Macrae et al., 2010). In Argonaute (Ago) proteins the PAZ domain anchors the 3' end of the miRNA. The Piwi domain of Ago-proteins has an RNase H fold and binds the miRNA at the 5'end in concert with the middle domain (MID) domain. Furthermore it harbors the catalytic DEDH motive, which is required for RNase activity (Meister et al., 2004; Nakanishi et al., 2012; Song et al., 2004; Yang and Steitz, 1995).

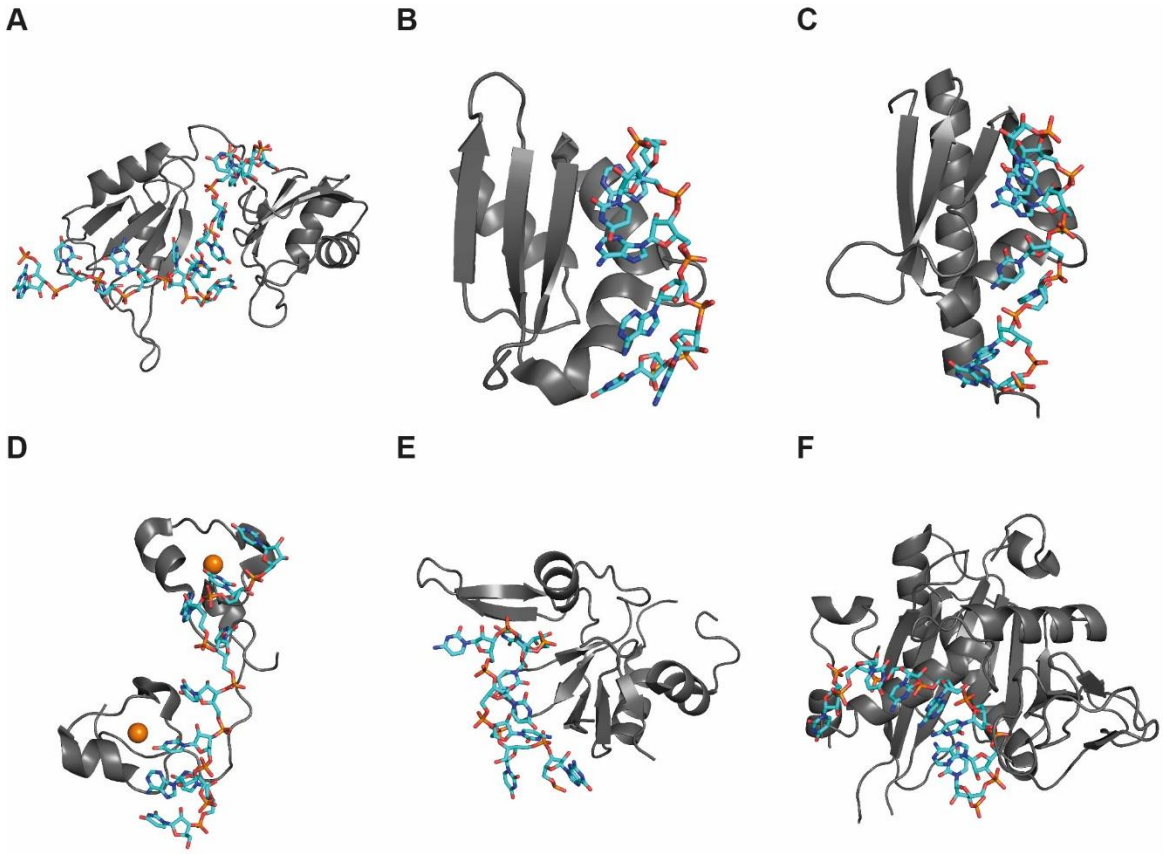


Figure 1.2: Overview of different single stranded RBDs in complex with RNA. (**A**) The tandem RRMs of Sex-lethal (PDB-ID: 1B7F). (**B**) Type I KH domain of Nova (PDB-ID: 1EC6) and (**C**) Type II KH domain of NusA (PDB-ID: 2ATW). (**D**) Two zinc fingers of TIS11d (PDB-ID: 1RGO). (**E**) The Ago2 PAZ and (**F**) Piwi domains (PDB-ID: 4W5N). For all single pictures, proteins are shown as grey ribbons, ssRNAs as cyan sticks and zinc ions as orange spheres.

1.3 The dsRBD fold

The dsRBD fold is present in many eukaryotic, prokaryotic and even viral RBPs. RBPs function in many biological processes, where recognition of dsRNA is required. In the following, some examples will be introduced, which make up only a marginal proportion of dsRBDs that were described in literature up to date. The *Drosophila* protein Staufen 1 (STAU1), for example, associates with three double-stranded stem-loop structures within the 3' UTR of *bicoid* mRNA and mediates the transport of the mRNA to the anterior pole of a freshly laid egg (Ferrandon et al., 1997, 1994). In humans, a complex of STAU1 and the long non-coding RNA (lncRNA) TINCR is required for somatic tissue differentiation (Kretz et al., 2013). Many family members of the adenosine deaminase acting on RNA (ADAR) enzymes comprise at least one dsRBD, which are followed by a deaminase domain. ADAR is able to specifically recognize and regulate dsRNA substrates, which are commonly found in 3' and 5' UTRs, intronic retrotransposons like Alu and long interspersed elements (LINEs) and also within miRNA precursors (pre-miRNA), by an A to I editing (Nishikura, 2006). DsRBDs play another important role in the processing of pre-miRNAs by Dicer and its partner proteins in a heterodimeric complex. In *Drosophila,* dmDcr-1

contains one dsRBD and its protein partner Loqs three dsRBDs (Bernstein et al., 2001; Förstemann et al., 2005).

The dsRBD fold was first identified in STAU1 from *Drosophila* and rbpa from *X. laevis*, where, on the basis of binding studies and computational analysis, a consensus sequence of 65-68 aas could be identified (St Johnston et al., 1992). Structural analysis of *E. coli* RNase III dsRBD and STAU1 dsRBD3 revealed a conserved fold (See Figure 1.3 A) containing α-helices and β-sheets with a αββββα topology (See Figure 1.3 B) (Bycroft et al., 1995; Kharrat et al., 1995). The N- and C-terminal α-helices (α1 and α2) stack against a core element consisting of three anti-parallel β-sheets (β1, β2, β3) through mainly hydrophobic interactions. Thereby, aliphatic side-chains are mainly found in the helices α1 and α2 and aromatic residues in the sheets β1 and β2 to complement the hydrophobic core. Structural data from complexes of dsRNA and dsRBPs indicate that the binding interface is distributed to three conserved regions from the dsRBD (Ryter and Schultz, 1998; Yang et al., 2010). Region 1 which is located in helix α1 consists mainly of a conserved glutamate. Region 2 is located in a loop region (L2) between β1 and β2 and contains a GPxH consensus motif that is required to anchor the dsRNA to the protein. Region 3 is located at the beginning of helix α2 and contains highly conserved positively charged residues forming the KKxAK consensus motif (Masliah et al., 2013). All three regions act simultaneously to specifically recognize an A-form dsRNA helix. E8 from Region 1 (Positions of the amino acids within the dsRBD were numbered according to the sequence alignment from Figure 1.3 A) thereby interacts with the minor grove of the dsRNA (See Figure 1.3 C). Region 3, with lysines K55, K56 and K58 from the KKxAK motif, specifically interacts with the major groove of the dsRNA and H31 from the GPxH motif of region 2 anchors the protein to the dsRNA by interacting with the succeeding minor groove.

Mutations in this regions strongly impact the affinity of the dsRBD to dsRNA and a triple mutant, where H31, K55 and K58 are mutated to Ala, even completely abolishes the affinity to a dsRNA substrate (Yamashita et al., 2011).

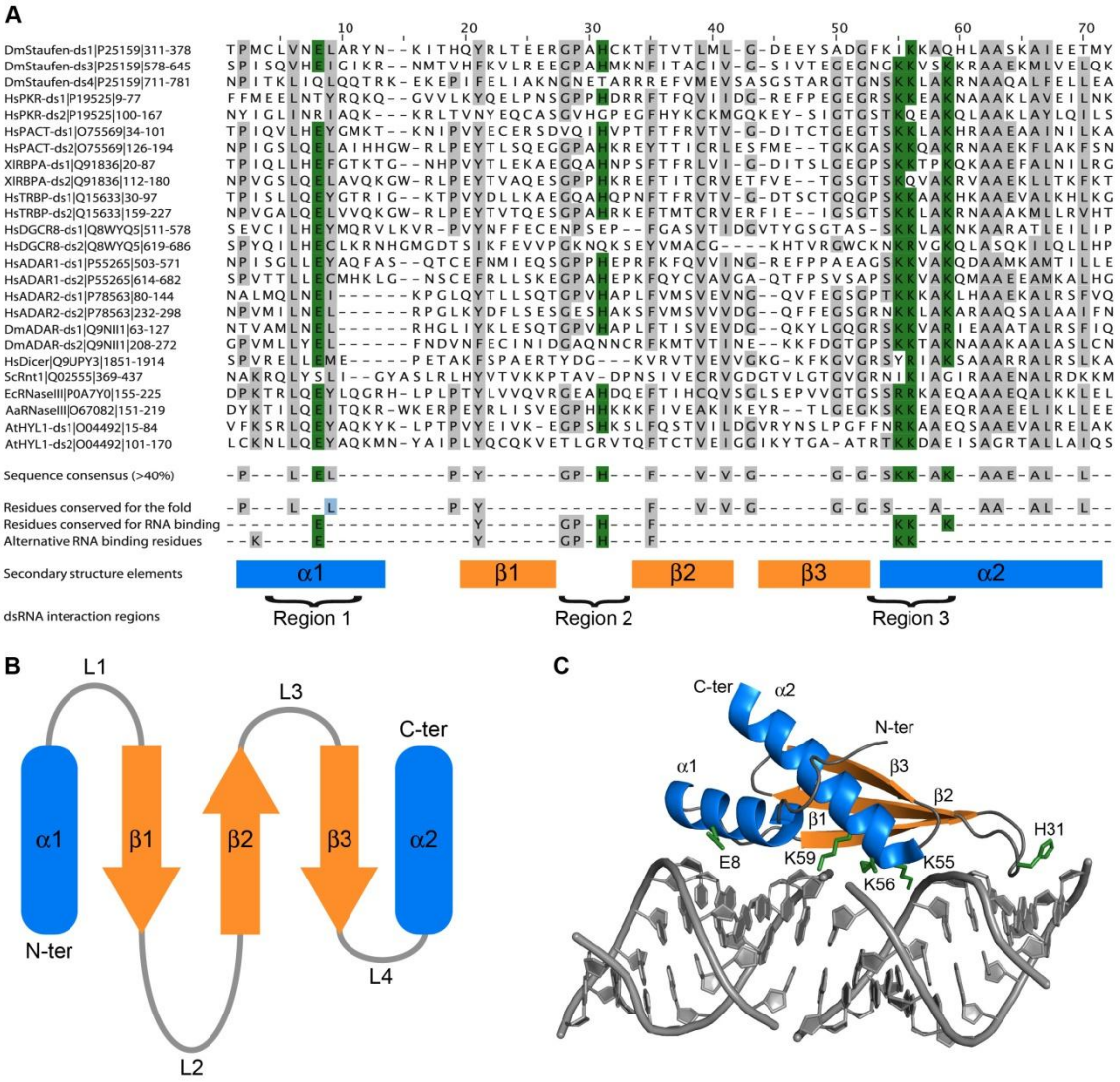


Figure 1.3: Characteristics of the dsRBD fold. (**A**) Sequence alignment of several dsRBDs from *Drosophila melanogaster* (Dm), *Homo sapiens* (Hs), *Xenopus leavis* (Xl), *Saccharomyces cerevisiae* (Sc), *Arabidopsis thaliana* (At), *Escherichia coli* (Ec) and *Aquifex aeolicus* (Aa). The names and positions of the dsRBDs and the accession code in the UniProt database (http://www.uniprot.org) are listed in front of the sequence. Important conserved residues for the folding and RNA binding are highlighted in grey, important residues from the RNA binding consensus motives in green. Adapted from Masliah et al., (2013). (**B**) Schematic overview of the dsRBD fold. Three antiparallel β-sheets (β1-3) are flanked by two helices (α1 and 2) and connected by four loops (L1-4). (**C**) Structure of the second dsRBD from TRBP in complex with dsRNA (PDB-ID: 3ADL). Important residues for the interaction are highlighted in green as in (**A**).

The dsRBD2 of *X. laevis* rbpa (Ryter and Schultz, 1998) constitutes the archetype of a canonical dsRBD with a sequence consensus of >40% from an alignment of several different dsRBDs (Masliah et al., 2013). However there also exist some dsRBDs which differ from the canonical sequence but still are able to bind dsRNA. The most variable part for a dsRBD is located within the N-terminus. The dsRBDs from *S. cerevisiae* RNase III, mammalian ADAR2 and *Drosophila* ADAR, for example, have shorter α1 helices, which has no detectable functional impact on the dsRBD function (Leulliot et al., 2004; Stefl et al., 2006). In case of *S. cerevisiae* RNase III,

however, there exists an additional helix α3, succeeding α2, which has been shown to have an impact on the stability of the domain and play an additional role in dsRNA binding. A second region of variability is the loop L2 between β1 and β2. In the *Arabidopsis* methyltransferase HUA ENHANCER 1 (HEN1) there are two dsRBDs with long insertions in L2. Crystallographic studies show, that this loop in the dsRBD1 binds to dsRNA and additionally provides protein-protein contacts with the methyltransferase domain, highlighting a bifunctional feature of the L2 insertion for both dsRNA binding and protein-protein interaction (Huang et al., 2009).

Interestingly several dsRBDs exist, which exhibit the classical dsRBD folding but instead of binding to dsRNA they are exclusively responsible for protein-protein interactions. Such an example can be found in the dsRBD3 of human Dicer interacting proteins TRBP and PACT, where the interface of the Dicer-TRBP complex has recently been solved by X-ray crystallography (Wilson et al., 2015). In the mammalian STAU1, which contains five dsRBDs, the dsRBD5 lacks residues and features required to bind duplex RNA. Together with a so called Staufen-swapping motif, STAU1 is able to form homodimers, which are functionally required for STAU1-mediated mRNA decay (Gleghorn et al., 2013).

Compared to the plethora of canonical dsRBD-structures, only little structural information of protein-protein interacting dsRBDs is available. Additional structural information of protein-interacting dsRBDs is required to elaborate the understanding of how dsRBPs are able to associate and collaborate with their binding-partners.

2. Small ncRNA mediated gene silencing

The discovery of the lin-4 gene locus in *Caenorhabditis elegans* (*C. elegans*) that does not encode for protein but express small RNA transcripts, which have a negative regulatory effect on translation of the lin-14 mRNA (Lee et al., 1993; Wightman et al., 1993), gave the headstone for the discovery of RNA interference (RNAi) by Fire et al. (1998). Although the effect of RNAi was already shown earlier in plants (Napoli et al., 1990), the finding that small ncRNA-mediated gene regulation is not only restricted to nematodes and plants but conserved among species attracted the interest of many researchers in this field and was finally honored with the Nobel Prize in 2006.

2.1 Classes of small ncRNAs

The development of deep sequencing methods identified a large number of small ncRNAs with different regulatory mechanisms. According to their structure, biogenesis and related function these small ncRNAs can be divided into three different groups, the miRNAs, the siRNAs and the piRNAs.

2.1.1 MiRNAs

MiRNAs are approximately 20-25 nts in length and are the predominant class of small RNAs in most somatic tissues. miRNA genes belong to one group of the most abundant gene families and are distributed among species and are even found in some viruses (Griffiths-Jones et al., 2008). In their canonical biogenesis pathway, miRNAs are derived from PolII transcribed primary miRNAs (pri-miRNA). However, there also exist miRNAs, which do not originate from canonical processing and where the mechanisms are not yet understood in detail. These miRNAs are generated from specialized introns (miRtrons), tRNAs, snoRNAs and capped transcripts (See Introduction Chapter 2.2 for detailed description of the miRNA biogenesis pathway).

When a mature miRNA is bound to one of its effector proteins from the Ago family, it functions as the so called miRNA-induced silencing complex (RISC), which is able to regulate fundamental biological processes like development, diff erentiation, apoptosis and proliferation. Therefore one strand of the miRNA duplex guides the RISC by Watson-Crick paring with its 5' nts 2-8, the so called "seed sequence" (Lewis et al., 2003), to a complementary region on the target mRNA leading to translational repression, destabilization and degradation of the respective target mRNA. Under consideration of all predicted miRNA binding sites that are conserved among species, at least 60 % of all protein-coding genes contain a potential binding site and adding also non-conserved sites it is likely that almost every protein-coding gene is under control of certain miRNAs (Friedman et al., 2009). Due to the relevance of the miRNA pathway, the biogenesis and

function of miRNAs is strictly regulated and dysfunction in regulation is often associated with various human diseases including cancer (Mavrakis et al., 2010) and neurodevelopmental disorders (Im and Kenny, 2012). The involvement of miRNAs as cellular key regulators offers the potential for the development of miRNA based therapeutics (Kota et al., 2009; Lanford et al., 2010; Melo et al., 2011), which underlines the importance for the examination of the molecular processes of miRNA biogenesis and regulation in every detail.

2.1.2 PiRNAs

PiRNAs are ~24-32 nucleotides in length and are specifically expressed in the germline. In complex with a member of the Piwi clade of Ago proteins, piRNAs are important for gametogenesis (Cox et al., 2000; Deng and Lin, 2002; Szakmary et al., 2005) and for silencing of mobile genetic elements such as transposons (Vagin et al., 2006) in many different species such as diverse as mammals, flies and fishes. The generation of piRNAs, however, differs strongly from other small RNA biogenesis pathways and is still not fully understood. Primary piRNAs (pri-piRNAs) are expressed as single strands and derive from sense and anti-sense transcription from discrete genomic loci (Brennecke et al., 2007), termed piRNA clusters, which do not function as a substrate for Drosha and/or Dicer (Das et al., 2008; Houwing et al., 2007; Vagin et al., 2006). In *Drosophila*, the generation of piRNAs requires two nucleases. First pri-piRNAs are processed by a protein called Zucchini (Pane et al., 2007), which like its mouse homologue PLD6 (Ipsaro et al., 2012) belongs to a member of the phospholipase-D family of phosphodiesterases. This family includes both, phospholipases and nucleases (Haase et al., 2010) and is supposed to be responsible for the generation of the 5' end from pri-piRNAs. After the 3' end of the pri-piRNA is trimmed to the right length (Kawaoka et al., 2011), the piRNA methyltransferase Pimet (Hen1) is required for 2′-O-methylation and final maturation of pri-piRNAs (Saito et al., 2007), which are further transferred to a Piwi protein. Then an amplification cycle takes place, the so called "ping-pong cycle". In the ping-pong cycle, mature sense pri-piRNAs guide Piwi proteins to complementary sequences on antisense transcripts from the respective piRNA cluster. As Piwi proteins possess a slicer activity, they are able to cleave the target antisense transcript from the cluster and generate a new 5′ end. This 5′ end is then bound again by another Piwi protein and the 3′ end is trimmed again to the length of a mature piRNA. This piRNA can now target again sense transcripts, transcribed from the piRNA cluster, leading to an enrichment efficient piRNAs against transposons. In *Drosophila*, the two Piwi proteins, which are responsible for secondary piRNA production, are called Aubergine and Ago3.

2.1.3 SiRNAs

SiRNAs comprise an average length of ~21 nucleotides. Similar to piRNAs, they exhibit a full complementarity to their target RNA, which leads to an endonucleolytic cleavage when incorporated into a slicer active Ago protein. Endogenous siRNA-production has been described first in *C. elegans* where primary siRNAs, which are derived from Dicer nuclease-mediated cleavage of the original long dsRNA trigger, and secondary siRNAs, which are additional small RNAs whose synthesis requires an RNA-directed RNA polymerase (RdRP), are generated. The generation of secondary siRNAs requires a specialized mechanism, which is fundamentally different from other conserved small RNA pathways (Pak and Fire, 2007; Yigit et al., 2006).

In *Drosophila*, there exists a small RNA pathway where Dicer 2 and its cofactors R2D2 and a specific splice isoform of Loqs (Loqs-PD) are involved in the generation of certain siRNAs from exo- or endogenous long dsRNA precursors (Hartig et al., 2009; Liu et al., 2003). Because of the relevance for this thesis, the mechanisms behind biogenesis from small non-coding regulatory RNAs in insects will be introduced in more detail in Chapter 2.3.

In mammalian somatic cells, there is no evidence for the existence of endo-siRNAs. On the one hand, a recombinant overexpressed and purified human Dicer is able to process long fully complementary dsRNA into ~21 nt long siRNAs in an ATP-independent manner (Zhang et al., 2002) *in vitro*, however, on the other hand, the transfection of mammalian cells with long dsRNA templates (38-1662 base pairs (bp)) does not result in efficient production of 21 nt long siRNAs (Caplen et al., 2000), but triggers interferon response by activation of protein kinase R (PKR) (Manche et al., 1992; Stark et al., 1998). Transfection of 21 nt long double stranded RNAs with a 2 nt overhang on the 3' end of each siRNA is able to overcome interferon response and leads to an effective knockdown of target genes in mammalian cells (Elbashir et al., 2001). Although containing the prerequisites to perform RNAi in general, mammalian somatic cells seem to react differently to long dsRNA species then shown in *C. elegans* or in *Drosophila*. Some exceptions, however, might exist in mouse oocytes and embryonic stem cells, where endo-siRNAs are generated upon viral infection, from transposable element-, bidirectional- or pseudogene-transcription or from long hairpin structures (Dautry et al., 2007; Tam et al., 2008; Watanabe et al., 2008).

2.2 Biogenesis of mammalian miRNAs

The biogenesis of mammalian miRNAs starts in the nucleus, where PolII transcribes long primary transcripts (~ 1kb), which contain at least one local hairpin structure, where a mature miRNA sequence is embedded. In humans, many of the canonical miRNAs are located in introns of coding and non-coding genes. Other miRNAs are organized in clusters in which several miRNA loci follow each other, forming a polycistronic transcription unit (Lee et al., 2002). Each

pri-miRNA consists of a stem of ~33-35 bp in length, a terminal loop and single stranded RNA sequences at the 5' and 3' ends. This specific structural feature is recognized by the nuclear RNase III enzyme Drosha, which interacts with its cofactor DiGeorge syndrome critical region 8 (DGCR8), the so called "microprocessor complex" (Figure 1.4, upper part) (Auyeung et al., 2013; Denli et al., 2004; Gregory et al., 2004; Han et al., 2004; Landthaler et al., 2004; Melo et al., 2011). The complex further catalyzes a precise cropping of the stem loop, where Drosha cleaves the hairpin approximately 11 bp away from the "basal" junction, the lower end of the pri-miRNA stem, and 22 bp away from the "apical" junction, the upper end of the pri-miRNA stem, resulting in a smaller ~65 nts long hairpin-shaped RNA, the so called precursor-miRNA (pre-miRNA) (Lee et al., 2003). A recently published crystal structure of human Drosha (Kwon et al., 2015) suggest that Drosha alone is capable of recognizing the key structural features in the basal side of pri-miRNA and measuring of the 11 bps for cleavage of the stem. DGCR8, on the other side, binds Drosha as a dimer and is required for the apical binding and positioning of the pri-miRNA. After positioning, one of the tandem RNase III domains (RIIID) cuts the 5' strand and the other one the 3' strand, generating a two-nucleotide-long 3' overhang forming the mature pre-miRNA (Han et al., 2006) (Figure 1.4, upper part).

After nuclear processing of the pri-miRNA, the nuclear shuttling protein exportin 5 (Exp5) in complex with its cofactor the GTP-binding nuclear protein Ran (RanGTP) binds to the pre-miRNA with its positive inner surface and forms the transport complex (Bohnsack et al., 2004; Lund et al., 2004; Okada et al., 2009; Yi et al., 2003). The tunnel like structure is able to strongly interact with the two-nucleotide-long 3' overhang of the pre-miRNA and enables Exp5 to recognize dsRNA stems of >14 bps in length. After accurate assembly of Exp5, RanGTP and the pre-miRNA, the complex is translocated from the nucleus to the cytoplasm through the nuclear pore complex (NPC). In the cytoplasm, GTP bound to RanGTP gets hydrolyzed, initiating a disassembly of the complex and releases the pre-miRNA.

In the cytoplasm (Figure 1.4, lower part), a second RNase III enzyme, called Dicer (Bernstein et al., 2001), binds to the pre-miRNA and cleaves the hairpin-structured RNA close to its terminal loop, liberating a small RNA duplex. As it is the case for Drosha, both Dicer RIIIDs form an intramolecular dimer, which creates the active center of the enzyme in a central cleft between both domains. The human Dicer anchors both the 3' and the phosphorylated 5' end within its PAZ domain and then recognizes the cleavage site by a predefined distance (~22 nucleotides), starting from the 5' end of the pre-miRNA by the so called 5' counting rule (Park et al., 2011; Tian et al., 2014). For hDcr it is shown that it is able to interact with the trans-activating response RNA-binding protein (TRBP) and the interferon-inducible double-stranded RNA-dependent protein kinase activator A (PACT). However, the molecular details of theses complexes were not understood yet in detail. TRBP is shown to have an impact on the cleavage sites of Dicer by stabilizing the protein/RNA complex and enhancing the catalytic turnover of the pre-miRNA.

This leads to a preferential processing of certain miRNA isoforms compared to the processing of hDcr alone (Fukunaga et al., 2012) (Figure 1.4, lower part).

The RISC loading step and the mechanistic details of miRNA-guided gene silencing are separately described in Chapters 2.4 and 2.5.

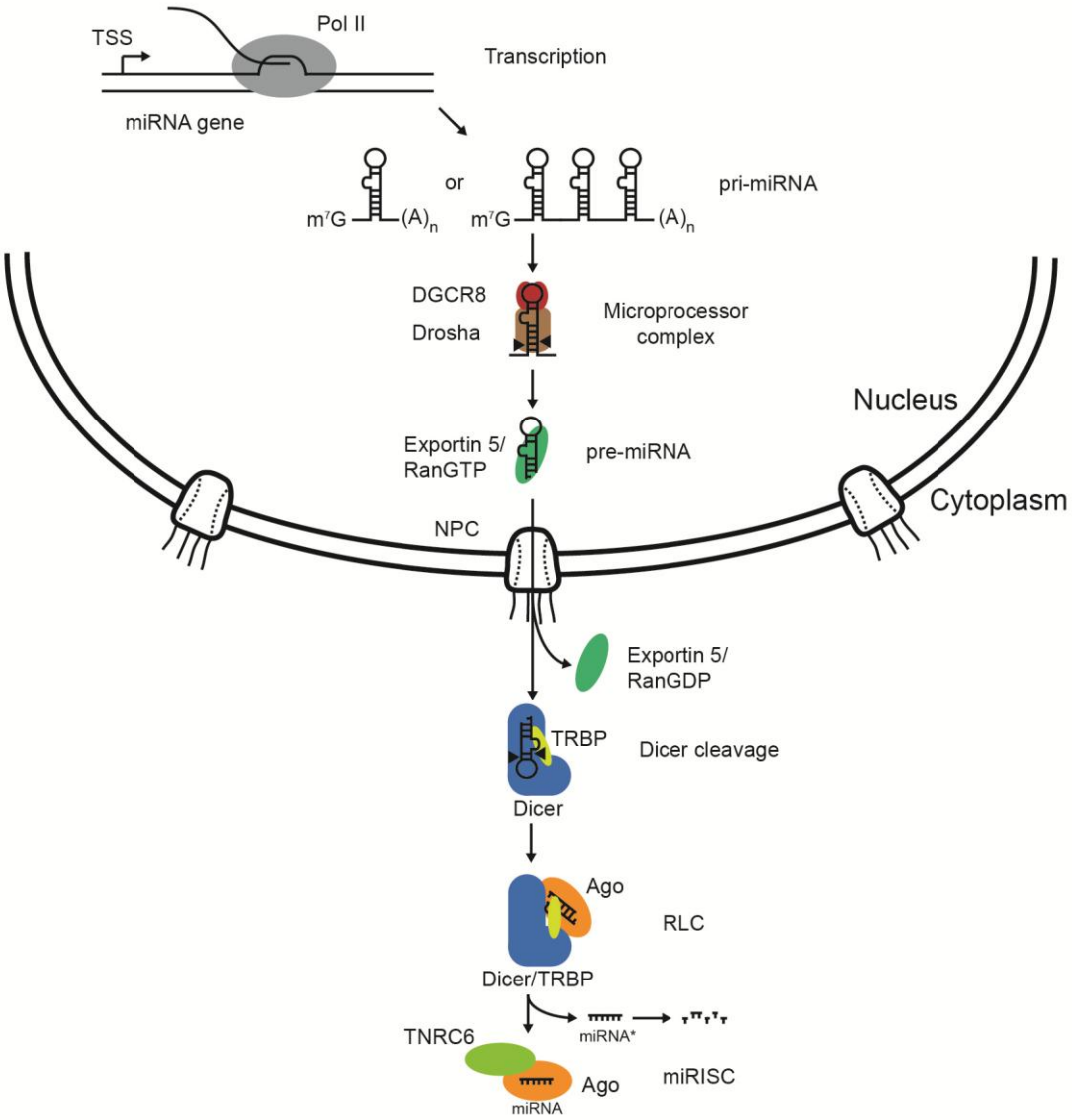


Figure 1.4: Biogenesis of miRNAs in mammals. PolII starts transcription of pri-miRNA from the transcriptional start site (TSS) of miRNA genes. Pri-miRNAs contain single or clustered stem loop structures within their sequences and, like mRNAs, were capped and polyadenylated. The microprocessor complex, containing DGCR8 and Drosha, binds to a single stem loop and releases the pre-miRNA which is recognized by a complex of RanGTP and Exp5 that initiates translocation through the nuclear pore complex (NPC) form the nucleus to the cytoplasm. After hydrolysis of GTP bound to RanGTP, the pre-miRNA is released from the complex and bound by Dicer and its cofactors TRBP or PACT. After removal of the stem loop, the mature miRNA is loaded onto an Ago protein (Ago1-4) where the guide strand (miRNA) gets incorporated and the passenger strand (miRNA*) is removed and degraded. In complex with a protein from the TNRC6 family, miRNA loaded Ago forms the active miRISC.

2.3 Biogenesis of small regulatory RNAs in insects

In contrast to the mammalian miRNA biogenesis machinery, where one single Dicer in complex with its dsRBP partner protein, is responsible to for the processing of miRNAs, there are two distinct isoforms of Dicer in the fruit fly *Drosophila*, dmDcr-1 and Dicer-2 (dmDcr-2). Both Dicer have distinct functions in the generation of small ncRNAs and were assisted by two different kinds of dsRBPs. DmDcr1 associates with Loqs and dmDcr-2 with R2D2 (Lee et al., 2004; Liu et al., 2003; Tomari et al., 2004). For the *loqs* gene, there are four different splicing variants (Hartig et al., 2009) described (*loqs RA, RB, RC* and *RD*) which are derived through alternative splicing and encode for functionally different proteins (Loqs-PA, -PB, -PC and -PD) that differ in their C-terminal region succeeding the first and second dsRBDs (Figure 1.5).

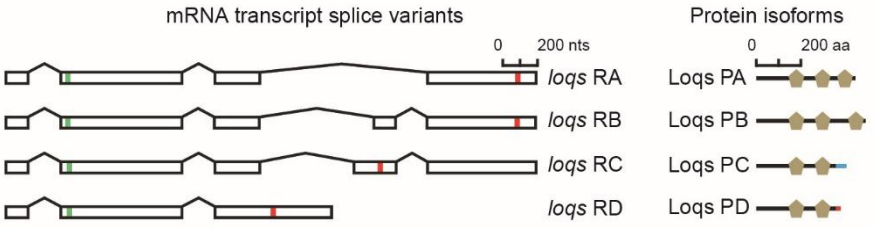


Figure 1.5: Loqs isoforms generated by alternative splicing. Left side: Different identified *loqs* mRNAs were depicted. Intronic sequences, which were removed upon splicing, were drawn as tin lines. Translational start sites were marked with green bars and stop codons were indicated with red bars. Right side: Protein sequence of different Loqs isoforms. Regions with a dsRBD fold were indicated as pentagons and alternative, unstructured and no dsRBD containing C-termini were highlighted in blue (Loqs-PC) and red (Loqs-PD). Figure adapted from (Hartig et al., 2009).

Loqs-PA and -PB both contain a third dsRBD, which allows the interaction with dmDcr-1 and the processing of pre-miRNA in the miRNA biogenesis pathway (Förstemann et al., 2005; Jiang et al., 2005; Saito et al., 2005; Ye et al., 2007). Both isoforms were expressed in *Drosophila* cells, but only the Loqs-PB seems to be necessary and sufficient for embryonic development and the miRNA pathway (Park et al., 2007). The miRNA biogenesis pathway from *Drosophila* is mostly conserved to the mammalian biogenesis pathway (Figure 1.4) and therefore will not be introduced in more detail. The only differences in flies are Pasha, which is equivalent to DGCR8 (Giot et al., 2003; Landthaler et al., 2004) and Loqs-PA and -PB, which are equivalent to TRBP and/or PACT.

Loqs-PC lacks the dsRBD3 that is present in Loqs-PA and -PB and only little is known about this isoform and if it plays a physiological role *in vivo* at all.

Isoform Loqs-PD, which has a different polyadenylation site in the third intron, contains an alternative 22 amino acid C-terminal sequence instead of the dsRBD3. Loqs-PD is highly expressed in *Drosophila* Schneider 2 (S2) cells but does not interact with the dmDcr-1 (Hartig et al., 2009). In earlier studies, proteins products from the *loqs* gene were thought to interact

exclusively with dmDcr-1. Therefore, it was surprising that a knockdown of the complete *loqs* gene also had an impact on the endogenous (endo)-siRNA processing pathway which, is made up mainly by dmDcr-2 and its dsRBP partner R2D2 (Czech et al., 2008). The dsRBP R2D2 consists of two dsRBDs ("R2") and a rather unstructured C-terminus that is able to interact with Dicer2 ("D2"). Loqs-PD, as R2D2, is able to interact with dmDcr-2 by its unique putatively unstructured C-terminal sequence and assists dmDcr-2 in the endo-siRNA pathway for a *de novo* response against invading selfish genetic elements in somatic cells. Both proteins have overlapping functions in targeting of transposons, however certain transposons display a preference for either dsRBP during production or loading (Mirkovic-Hösle and Förstemann, 2014).

Besides its function in the endo-siRNA pathway, it has been shown, that R2D2 together with dmDcr-2 is able to direct innate immunity against viruses (Wang et al., 2006). The separation of the dicing processes into two different pathways might therefore result from a strategy, which allows a rapid counteracting to invading viral long dsRNA, which would, in the case of one Dicer, compete with the endogenous pre-miRNAs for processing and loading to an Ago protein (Ghildiyal and Zamore, 2009). It is also likely that an evolutionary pressure from rapidly evolving viral strategies to escape the RNAi-pathway lead to the specialization of dmDcr-2 and Ago2 which is supported by the finding that *dcr-2* and *ago2* belong to the most rapidly evolving *Drosophila* genes (Obbard et al., 2006).

In mammals, on the other side, the RNAi pathway for counteracting viral infection might have been substituted by the evolution of an elaborate, protein-based immune system (Vilcek, 2006; Williams, 1999). Notably, it has been shown that in mouse embryonic stem cells (mESCs), there is a RNAi response upon infection with an encephalomyocarditis virus (EMCV) or Nodamura virus (NoV), as ~22-nucleotide RNAs derived from viral dsRNA replication intermediates accumulate in the cell and were incorporated into Ago2 (Maillard et al., 2013), suggesting a possible active, Dicer-dependent, antiviral RNAi response in mammalian cells or at least in embryonic stem cells that lack an IFN response (Billy et al., 2001; Paddison et al., 2002).

2.4 RISC loading

After Dicer processing a dsRNA of ~22 nts with 2 nt 3' overhangs on each side of the duplex is generated. The mature siRNA or miRNA is further transferred to a member of the Ago clade. Thereby, the dsRBP partner of Dicer and the thermodynamic properties of the duplex itself determine, which strand gets incorporated into the Ago protein, the so called "guide strand" (miRNA) and which strand is released and gets degraded, the so called "passenger strand" (miRNA*) (Schwarz et al., 2003).

2.4.1 RISC loading in *Drosophila*

In *Drosophila* the molecular details of Ago loading were most extensively studied yet and many components involved in this process have been identified. Due to the existence of two Dicer, also RISC loading is performed separately, where miRNAs associate in miRISC and siRNAs in siRISC (Kawamata et al., 2009). For siRNA-loading, a RISC loading complex (RLC) consisting of dmDcr-2, R2D2, Ago2 and several auxiliary proteins is assumed. The nuclear TATA-binding protein-associated factor 11 (TAF11) has been identified by a genetic screen to be involved in siRISC formation. The mainly nuclear protein co-localizes with dmDcr-2/R2D2 in so called D2 processing bodies in the cytoplasm. There a TAF11 tetramer facilitates dmDcr-2/R2D2 tetramerization and enhances siRNA binding and RISC loading activities (Liang et al., 2015). In several studies it has been shown that Ago loading requires ATP and the help of a complex of chaperones which also can be observed in animals and plants (Iki et al., 2010; Iwasaki et al., 2010). A recent study identified the proteins Ago2, dmDcr-2, R2D2, Hsc70, Hsp90, Hop, Droj2 (an Hsp40 homologue) and p23 to be sufficient to reconstitute RISC loading *in vitro* (Iwasaki et al., 2015). Omitting any component leads to a deficiency of Ago loading. The chaperone complex is required to hold Ago2 in a productive state and extending the dwell time of the dmDcr-2/R2D2/siRNA complex on Ago2 for accurate 5' phosphate recognition of the siRNA guide strand. Upon loading, R2D2 is required for sensing the siRNA duplex for thermodynamic asymmetry and promoting an orientation, where the 5' phosphate of the guide strand is bound by Ago2 (Tomari et al., 2004). After loading, subsequent unwinding and passenger stand ejection has been shown to be ATP and chaperone independent for *Drosophila* Ago2 (Iwasaki et al., 2010; Kawamata et al., 2009; Yoda et al., 2010). DmDcr-2 in complex with R2D2 acts as a gatekeeper for the assembly of Ago2 complexes, where the loading of Ago2 with fully complementary siRNAs is favored (Tomari et al., 2007).

For the assembly of miRISC, on the other hand, little is known. Efficient Ago1 loading is dependent on the occurrence of central and terminal mismatches and a 5'-Uracil, which are a typical structural features of most miRNA/miRNA* duplexes (Förstemann et al., 2007; Kawamata et al., 2009; Seitz et al., 2011). However, if dmDcr-1 and Loqs are also required for the miRNA loading process or if DmDcr-2 and R2D2 might additionally play a role within this process is not fully understood.

2.4.2 RISC loading in mammals

For mammalian RISC loading, many studies have been performed, to investigate the molecular details of the Ago loading step. However, many aspects are still not elucidated in detail and some are even contradictory. Ago2 is capable to bind to the miRNA-generating hDcr/TRBP complex (Chendrimada et al., 2005; Haase et al., 2005; Meister et al., 2005; Mourelatos et al., 2002).

Consistently, a ternary complex of Ago2, hDcr and TRBP can be associated *in vitro* and the structure could be resolved by cryo-electron microscopy (Wang et al., 2009). For a siRNA duplex, it is shown that, after processing of the dsRNA substrate, a repositioning along the helicase domain of the mature siRNA duplex takes place within the hDcr/TRPB complex and the strand is scanned for the thermodynamic less stable end by a putative second binding site in hDcr (Noland et al., 2011). By the biased arrangement of the dsRNA duplex a strand-selective RISC loading of Ago is accomplished. Additionally, it has been shown in a Dicer-knockout rescue assay with a dsRBP-binding deficient mutant of hDcr, that for a subset of 108 analyzed miRNAs there is a pronounced change in the proportion of guide versus passenger strand incorporation into Ago2 (Wilson et al., 2015). However, there also exist sequencing-data from another miRNA screen in a TRBP and PACT knockout background, where no difference in strand selection and Ago loading could be observed in the absence of the Dicer partner proteins (Kim et al., 2014). Both studies commonly agree with the finding that in the absence of the Dicer partner proteins, the levels of several iso-miRNAs were altered. An iso-miRNA is derived from a mature miRNA, which can be altered by substitutions, insertions or deletions, 3' end non-templated additions, and 5' and/or 3' cleavage variations (Morin et al., 2008). Therefore, dsRBPs seem to have a relevant meaning on the positioning of the pre-miRNAs upon Dicer-cleavage, which on the other hand determines the length of the mature miRNA.

As in flies, an effective loading and incorporation of small ncRNA into an Ago protein requires a heat shock protein. Hsp90 has been identified to stabilize Ago2 in an open state, allowing an accurate loading of the RNA duplex (Johnston et al., 2010). Hsp90 thereby is assisted by the co-chaperones FKBP4 and FKBP5. A depletion of either co-chaperone leads to a strong decrease of Ago2 levels due to the instability of unloaded Ago within the cell (Frohn et al., 2012; Martinez et al., 2013). The chaperone complex keeps Ago in a loading-competent state until a mature miRNA is bound.

Once bound to Ago, the guide strand gets incorporated and the passenger strand is removed by an unwinding of the duplex. The N-terminus of Ago itself has been suggested to contribute to the unwinding process through an active wedging process (Kwak and Tomari, 2012). Perfect complementarity in the middle of siRNAs and miRNAs allows catalytic active Ago proteins to position the RNA duplex within their catalytic DEDH motive, which activates the slicer activity and nicks the passenger strand. The endonuclease C3PO has been identified to interact with Ago2 and degrades the Ago2-nicked passenger strand, contributing to an accelerated RISC formation (Ye et al., 2011; Zhang et al., 2016). Catalytic inactive Ago proteins, on the other side, were also efficiently loaded but the removal of the passenger strand is kinetically less favored (Dueck and Meister, 2014; Meister, 2013). Once incorporated into an Ago protein, the guide strand is tightly bound and cannot be exchanged between different members of the Ago protein family (Dueck et al., 2012).

2.5 Mechanisms of miRNA-mediated gene silencing

Once incorporated into Ago, the miRNA guide strand directs Ago to its seed-complementary target mRNA allowing PTGS. To be able to mediate PTGS, a loaded Ago-protein binds to an effector protein from the GW182 protein family. GW182 proteins are mainly unstructured and contain multiple Trp residues, which are often flanked by Gly (GW repeats). *Drosophila* possesses one family member namely GW182 and mammals possess three paralogs namely trinucleotide repeat-containing protein 6 A, B, and C (TNRC6 A, B and C) (Baillat and Shiekhattar, 2009; Meister et al., 2005; Pfaff and Meister, 2013). These proteins are also commonly referred to as GW proteins.

GW proteins comprise a modular composition and are able to act as scaffolds for a plethora of different regulatory proteins. The N-terminal part of GW proteins is Trp-rich and serves as an interaction platform for Ago proteins, the so called "Ago-hook" (Chekulaeva et al., 2009; Lian et al., 2009; Takimoto et al., 2009; Till et al., 2007). Thereby, two Trp residues with an exact spacing between each other were required, to specifically perform an high-affinity interaction with a hydrophobic pocket in the Piwi domain of Ago proteins (Hauptmann et al., 2015; Pfaff et al., 2013; Schirle and MacRae, 2012). The C-terminal silencing domain comprises a ubiquitin-associated (UBA)-like domain, a glutamine (Q)-rich domain, a poly(A)-binding protein interacting motif 2 (PAM2) and an RNA-recognition motif (RRM) and is able to silence bound transcripts independently of Ago proteins (Lazzaretti et al., 2009). In this C-terminal part of the protein, Trp-based interactions with the deadenylation complexes CCR4-NOT and PAN2-PAN3 take place (Chekulaeva et al., 2011; Christie et al., 2013; Fabian et al., 2012). A removal of the poly(A) tail, which is approximately 50–100 adenosines in length (Chang et al., 2014; Subtelny et al., 2014), by the deadenylation complex causes an initial destabilization of the mRNA. This destabilization is further enforced by the attraction of the decapping complex, containing the enhancer of decapping 3 (EDC3), EDC4, decapping protein 1 (DCP1) and DCP2 by the DEAD box protein 6 (DDX), which interacts with the decapping factors and the CCR4-NOT complex (Chen et al., 2014; Mathys et al., 2014; Tritschler et al., 2009). Once the 5' cap is removed, the truncated mRNA is recognized by the 5' to 3' exoribonuclease 1 (XRN1), which leads to the final degradation of the complete mRNA. The degradation of miRNA targets accounts for most of the miRNA-mediated repression at steady state levels in cultured mammalian cells (Jonas and Izaurralde, 2015).

3. The TRIM family

The introduced miRNA biogenesis pathway is an example for a highly specialized cellular process that is responsible for small RNA mediated gene silencing. All components within this pathway have very specialized and limited functions. Proteins from the TRIM family, however, are involved in diverse cellular processes, including apoptosis, cell cycle regulation, neurogenesis, muscular physiology and innate immune responses. To establish this functional diversity, TRIM proteins possess a variety of different protein domains, where even some have been identified to be direct RBDs (Loedige et al., 2014). In the following, the main characteristics, which assign proteins to the TRIM family and their respective functional subclasses, will be introduced.

The family of TRIM proteins can be characterized by the occurrence of a its eponymous characteristic motif, which consists of a "really interesting new gene" (RING) finger domain, at least one, in most cases two, B-Box-type zinc fingers (BB1 and BB2) and a Coiled-Coil (CC) domain. The TRIM motive is always located at the N-terminal part of the protein and the arrangement and the spacing between the individual domains is highly conserved among the family members (Reymond et al., 2001).

The RING finger domain is an ubiquitin ligase (also referred to as E3). E3 ligases catalyze the attachment of polyubiquitin chains to their respective substrates. The domain architecture consists of a series of cysteine (Cys) and histidine (His) residues, in the order of Cys-X_2-Cys-X_{9-39}-Cys-X_{1-3}-His-X_{2-3}-Cys-X_2-Cys-X_{4-48}-Cys-X_2-Cys, where "X" represents any amino acid. Cys in position 1, 2, 5 and 6 coordinate the first zinc ion and His 4 and Cys 3, 7 and 8 the second one (Barlow et al., 1994; Borden et al., 1995; Freemont, 1993). This cross brace arrangement, with the two zinc atoms coordinated by the conserved Cys and His residues at the center of the domain, is required for the domain architecture and for its E3 ubiquitin ligase (Joazeiro and Weissman, 2000). A mutation of a single Pro residue is sufficient to disrupt the ubiquitin transfer activity, as it is the case in the nematode protein Lin-41 (Budhidarmo et al., 2012; Tocchini et al., 2014). In certain uncommon TRIM family members, like the *Drosophila* proteins Brat and Wech, the RING domain is even missing.

The B-box (BBs) domains also contain the structural element of cross brace coordinated zinc atoms that are also present in the RING, ZZ and U-box domains of E3 and E4 ubiquitin ligases (Massiah et al., 2007; Tao et al., 2008). There are two types of BB (BB1 and BB2), where BB1 has the common motif Cys-X_2-Cys-X_{6-17}-Cys-X_2-Cys-X_{4-8}-Cys-X_{2-3}-Cys/His-X_{3-4}-His-X_{5-10}-His and, if present, always is located N-terminal from BB2 with the motif Cys-X_{2-4}-His-X_{7-10}-Cys-X_{1-4}-Asp/Cys-X_{4-7}-Cys-X_2-Cys-X_{3-6}-His-X_{2-5}-His. Altogether there is little known about the function of BBs. The structural similarity of BB to the RING domains might support the hypothesis that these domains are able to act as E3 ubiquitin ligases by their own, or promote the E3 activity of the RING domain. This is supported by the finding that the B-boxes of TRIM18/MID1 are mono-

ubiquitylated and B-box1 is able to enforce the E3 activity of the RING domain (Han et al., 2011).

The third part of the tripartite motif is the coiled-coil (CC) domain, which follows BB2 in all TRIM proteins (Reymond et al., 2001). The CC domain contains roughly a hundred, non-conserved amino acids that form larger α-helical secondary structures, which are intertwined to a rope like structure, stabilized by hydrophobic contacts between helices, which are often mediated by leucines (Leu) (Woolfson and Alber, 1995). The CC domain allows the formation of homo- or heterodimers between TRIM proteins, promotes the formation of high molecular weight protein complexes and defines subcellular TRIM-specific compartments (Reymond et al., 2001; Sanchez et al., 2014). The dimer architecture is conserved among several TRIM family members.

The occurrence of TRIM at the N-terminus dedicates a protein to the TRIM family. The functionality among certain family members however is achieved by a strong diversity of functional domains within the C-terminal part of the protein. Therefore TRIM proteins are assigned to certain classes according to their C-terminus (Short and Cox, 2006).

3.1 Classes of TRIM proteins

A systematic arrangement of TRIM proteins was established by Short and Cox, (2006) where they used a bioinformatical approach for the sub-classification of the entire human TRIM ensemble, based on their variety within the C-terminal domain compositions. So far, 10 different C-terminal domains have been identified, which succeed the TRIM either alone or in combination of each other. Under consideration of the subcellular localization, the expression levels in different tissues and the respective functions, TRIM proteins can be divided into nine different subclasses (C-I to C-IX). Figure 1.6 shows an overview of different TRIM protein family members and their classification according to the modular arrangement of their TRIM and C-terminal domain(s).

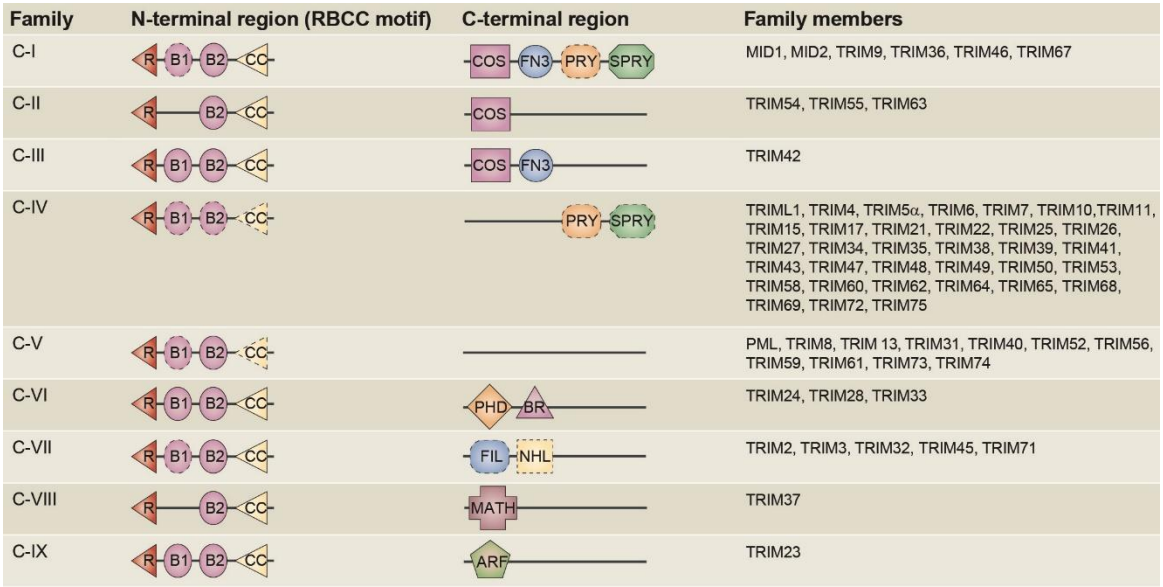


Figure 1.6: Schematic depiction of human TRIM protein family members. The classification of the TRIM families, depending on different carboxy-terminal-domain composition, was performed according to bioinformatics analysis of Short and Cox, (2006). Dotted lines indicate a variability of domain positions within the sequence. Domains are abbreviated in the following manner ARF: ADP ribosylation factor-like, BR: bromodomain, COS: C-terminal subgroup one signature, FN3: fibronectin type 3, FIL: filamin-type immunoglobulin, MATH: meprin and tumour-necrosis factor receptor-associated factor homology, MID: midline, PHD: plant homeodomain, PML: promyelocytic leukaemia. Adapted and modified from (Ozato et al., 2008).

The C-terminal subgroup one signature (COS) was found to be present in subgroups C-I, C-II and -III. The COS domain is located directly after the CC sequence in a subset of TRIM proteins and could also be found in non-TRIM proteins containing a CC domain (Short and Cox, 2006). The COS domain is required for localization of the respective TRIM proteins with microtubules (Short and Cox, 2006) and plays a role upon retroviral infection (Uchil et al., 2008; Yap et al., 2004).

The PRY domain with an average length of ~61 amino acids and the SPRY domain with ~140 amino acids (Figure 1.6, subclass C-IV) are the most common protein domains among the TRIM family, where the SPRY domain alone can be found in a total number of 39 family members. Due

to their high conservation even in plants and fungi, the SPRY domain is thought to be evolutionary ancient. In 24 human TRIM family members, however, the SPRY domain is fused to a PRY domain, forming a PRYSPRY domain (also called B30.2). This fusion is derived from a more recent evolution by the formation of adaptive immune mechanisms and therefore is restricted to vertebrate species including humans, mice, chickens and frogs (Rhodes et al., 2005; Ruby et al., 2005). Structural data of the PRYSPRY domains revealed a rigid binding platform to mediate protein-protein interaction between donor and acceptor sequences, which resembles interactions reported for e.g. antigen-antibody contacts (Grütter et al., 2006; Woo et al., 2006). Additional structural work on the PRYSPRY domain of the autoantigen TRIM21 revealed a canonical binding interface comprised of two discrete pockets formed by antibody-like variable loops, which are able to specifically interact with the heavy chain of IgG (James et al., 2007; Keeble et al., 2008). For TRIM20 (also called pyrin) it is shown that mutations within its PRYSPRY domain affect the interaction with pro-interleukin-1β (pro-IL-1β), which leads to increased inflammation rates in the hereditary disease "familial Mediterranean fever (FMF)" (Aksentijevich et al., 1997; Weinert et al., 2015, 2009).

The NHL-domain (Figure 1.6, subclass C-VII) was first described in the TRIM proteins NCL-1/HT2A/LIN-41 and consists of a five- or six-bladed β-propeller, which is arranged in a barrel-like assembly with a solvent filled cavity in the middle (Slack and Ruvkun, 1998). The NHL domain forms a rigid binding platform to meditate either protein-protein or protein-nucleic acid contacts. The NHL domain of TRIM32 for example is able to interact with the stress response factor p53, leading to apoptosis by its degradation (Liu et al., 2014). Consistently, an overexpression of TRIM32 strongly promotes oncogenic transformation and tumorigenesis in mice and it is also observed to be frequently overexpressed in different types of human tumors. TRIM3, on the other hand, has been identified as *bona fide* tumor suppressor, where its RING- and NHL-domains, both were required to promote ubiquitination of p21, which is required for the proliferation of a subset of glioma cells (Raheja et al., 2014). Several TRIM-NHL proteins have been identified to interact with Ago proteins and other miRNP components with diverse functional implications. Mammalian TRIM32 and *C. elegans* NHL-2 are able to enforce miRNA-mediated repression (Hammell et al., 2009; Schwamborn et al., 2009), whereas mammalian TRIM71 and *Drosophila* Mei-P26 negatively influence the miRNA biogenesis pathway (Neumüller et al., 2008; Rybak et al., 2009). Some NHL domains, however, show an accumulation of positively charged amino acids on the top surface of the β-propeller, whereas the bottom side is mostly negatively charged (Loedige et al., 2014). TRIM71 has been suggested to function as a direct translational regulator of many mRNA targets either independently or cooperatively with Ago proteins (Loedige et al., 2012). Consistently, TRIM71 has been found to directly interact with RNA in mouse embryonic stem cells (Kwon et al., 2013).

The filamin domain has a immunoglobulin-like structure with seven β-strands oriented in two antiparallel β-sheets (Bork et al., 1994) and is often associated with the NHL repeats at the C-terminus of class C-VII TRIM proteins.

The remaining C-terminal domains are the PHD domain, which is always in combination with a bromodomain (Figure 1.6, subclass C-VI) and supposed to be involved in chromatin-mediated transcriptional regulation (Friedman et al., 1996; Ivanov et al., 2007), the FN3 domain, which can contain binding sites for DNA and heparin, the ARF domain (Figure 1.6, subclass C-IX), which is involved in intracellular and vesicular trafficking and finally the MATH domain (Figure 1.6, subclass C-VIII), which is necessary and sufficient for self-association and receptor interactions by some TRAF proteins (Ozato et al., 2008).

3.2 The *Drosophila* TRIM-NHL protein Brat

Brat was first identified in a screen for lethal mutations in the *dopa decarboxylase* region (Wright et al., 1976) and was named after the phenotype, derived from inactivation of both alleles which leads to the production of a tumor-like neoplasm in the larval brain of *Drosophila* (Arama et al., 2000). Brat contains a NHL-domain and is an atypical representative of the TRIM family because it only contains two BB domains and the CC domain and lacks the RING domain. The NHL-domain contains a six bladed β-propeller, where each blade is composed of a highly twisted four-stranded antiparallel β-sheet, which is similar to the WD40 fold (Edwards et al., 2003). Brat has been shown to have an regulatory influence on the segmentation gene *hunchback* (*hb*), which is a key gene for the early pattern formation process in the Drosophila embryo (Lehmann and Nüsslein-Volhard, 1987; Tautz et al., 1987). For the maternally provided *hb* mRNA, which is distributed equally in the embryo, this translational repression is active until specific translation of the mRNA at the anterior pole of the embryo, which defines the body axis and directs thorax head formation (Hülskamp et al., 1990; Tautz and Pfeifle, 1989). A dysregulated expression of the Hb protein at the posterior pole of the embryo causes deficiencies in the abdominal segmentation. Additionally, The bottom face of the NHL domain of Brat was identified to constitute an interaction platform for the cap binding protein d4EHP, which contributes to the inhibition of *hb* mRNA translation (Cho et al., 2006).

Beside its function in the larval abdominal development of *Drosophila* early embryos, Brat has been identified to play a crucial role in the asymmetric cell division of neuronal stem cells (Brand and Livesey, 2011; Gómez-López et al., 2014; Knoblich, 2010). These stem cells are called neuroblasts (NBs) and the mechanisms underlying the processes of asymmetric stem-cell division are conserved among species. Upon the asymmetric cell division, the stem cell divides into a self-renewing NB and an intermediate neuronal progenitor (INP) cell, which finally differentiates after several few more cell divisions. To be able to initiate differentiation, the INP needs to suppress

self-renewal processes. This is achieved during asymmetric cell division by segregation of the cell fate determinants Brat, Numb and Prospero to the differentiating daughter cell (Betschinger et al., 2006; Knoblich et al., 1995; Rhyu et al., 1994). A disturbance of this process, where Brat is missing in the INP cell, causes an uncontrolled expansion of NBs and promotes tumor formation (Betschinger et al., 2006; Lee et al., 2006). In Brat mutant flies, the Brat targets Deadpan (Dpn) and Klumpfuss (Klu), which are involved in the self-renewal machinery, were strikingly increased in expression and lead to increased tumor formation. This tumorigenic phenotype, however, can be suppressed upon removal of either *dpn* or *klu* (Janssens and Lee, 2014; Xiao et al., 2012). In a genome-wide microarray expression study from *brat* homozygous mutant flies, which show a strong neoplastic adult brain phenotype, a large number of genes were identified, which showed highly significant changes in expression levels (Loop et al., 2004). Besides genes, which are relevant for asymmetric NB division, many genes involved in ribosome biogenesis, translation, and RNA processing could be identified, indicating also a contribution of these genes to the observed uncontrolled, tumor-like cell growth. Consistently, for the functional homolog of Brat in *C. elegans*, namely NCL-1, it is shown, that a loss of NCL-1 resulted in increased size of nucleoli, indicating higher levels of rRNA that promote ribosome biogenesis. This phenotype could be rescued by expressing Brat in the respective NCL-1 mutant worms (Frank et al., 2002; Frank and Roth, 1998).

Coming back to *hb*-regulation and early development in *Drosophila*, it was initially thought that the core functional component for translational repression of target mRNAs was Pum, that binds to its 8-nucleotide consensus motive UGUANAUA, where N stands for any nucleotide (Gerber et al., 2006; White et al., 2001) and recruits Brat and Nanos (Nos) via protein-protein interactions (Sonoda and Wharton, 2001, 1999). However, it has been shown that Brat binds directly to the *hb* mRNA in a Pum-independent manner (Loedige et al., 2014), suggesting that there are targets, which could be regulated by Brat also in the absence of Pum. Recently, it has been shown that the Brat dependent repression of *src64B* mRNA, which is important for axon maintenance, is independent of Pum (Marchetti et al., 2014). The discovery of a Pum independent binding of Brat to several targets mRNAs raised the question, whether Brat possesses a sequence-specificity to recognize its own targets. In a genome-wide analysis of mRNAs associated with Brat, a RNA motif with the consensus NNUGUUDNN (D=A/G/U) could be identified (Laver et al., 2015). Brat and Pum associate with hundreds of mRNAs in early embryos, where only one third are co-bound by both proteins. The occurrence of the Brat motif in proximity of the Pum motif lead to the assumption that Brat is able to interact with Pum by protein-protein interaction to allow a formation of a quaternary complex of Brat/Pum/Nos and the *hb* target mRNA (Edwards et al., 2003; Sonoda and Wharton, 2001). However, it has been shown in gel filtration experiments that the isolated Brat-NHL and the Pumilio homology domains (Pum-HD), that is responsible for RNA binding (Wang et al., 2001), formed a stable complex only in the presence of RNA (Loedige

et al., 2014). Therefore the Pum motif might be required for an altered, more accessible RNA structure that would facilitate binding of the Brat-NHL domain. The association of either Pum or Brat alone or the simultaneous binding of both proteins to a target mRNA influences the translational status and the stability of the mRNA (Laver et al., 2015).

For Pum, the molecular details for the interaction its Pum HD with a target RNA substrate are supported by several structural studies (Edwards et al., 2001; Gupta et al., 2008; Wang et al., 2002, 2001), explaining the observed sequence-specificity in an atomic detail. Brat on the other side, has only recently been discovered to directly bind to RNA (Kwon et al., 2013; Loedige et al., 2014) and the structural details for this interaction were still unknown when this PhD project was started.

Aims of this thesis

RBPs play a central role in post transcriptional gene regulation mechanisms that act on the level of mRNA transcripts. The mechanisms of how RBPs are able to interact with their target sequences are diverse and differ among RBP families.

Ago proteins associate with small ncRNAs that are generated by a complex cellular processing machinery where several proteins including Dicer and certain dsRBPs are involved. After loading, the small RNA provides Ago with "sequence-specificity" to identify distinct target RNAs. Thereby, some Ago proteins are endonucleolytically active and are able to cleave their targets, while others lack this feature. Ago proteins are able to recruit additional factors that mediate translational repression and target destabilization.

The aim of the first part of this thesis was to investigate the mechanisms behind the cytosolic processing of miRNAs and their loading into Ago proteins. The cytosolic processing of pre-miRNAs is performed by a complex of the RNase III enzyme Dicer and a dsRBP partner. Using structural biology and functional analysis, this thesis aimed to identify the molecular basis of the interaction platform between dmDcr1 and the dsRBD3 of Loqs and the implications on the processing of miRNAs. Furthermore an assay should be generated to be able to observe the downstream process of Ago loading with small fluorescently labeled ncRNAs using FCCS.

TRIM-NHL proteins are conserved among multicellular eukaryotic organisms and control cell fate decisions in various stem cell linages. *Drosophila* Brat has been identified as an RBP that is able to repress translation of several mRNAs upon directly binding to a consensus motif. The aim of the second part of this thesis was to structurally characterize the molecular details of the interaction between the NHL domain of Brat and a consensus motif RNA target sequence to obtain information of how the sequence specificity is brought about.

II. Results

1. Investigation of the human RISC loading complex

1.1 FCCS-based Argonaute (Ago) loading assay

During RISC loading small RNAs are generated by Dicer and transferred to a member of the Ago protein family. For the characterization of the binding of a small double stranded RNA (dsRNA) substrate to an Ago protein, a dual-color fluorescence cross-correlation spectroscopy (FCCS) assay was designed. FCCS is a confocal two-color spectroscopic method, where two laser lines were used to illuminate a microscopic detection volume in the range of one femtoliter (Bacia and Schwille, 2007). Fluorescently labeled molecules enter or leave this volume by diffusion and thereby generate a fluctuation in their fluorescence. From the fluctuation curves of each labeled component, various biophysical data such as molecular mobility, concentration and equilibrium and rate constants of molecular interactions can be extracted. To investigate Ago loading processes, green fluorescent protein (GFP) was fused N-terminally to Ago and the siRNA was labeled on its 3'end with a Cy5 fluorophore. After the initial characterization of the binding of one labeled siRNA to Ago, the affinity of any unlabeled siRNA can be determined in a competition reaction (see Figure 2.1 for schematic representation of the assay).

To validate the hypothesis that knockdown efficiencies of siRNAs correlate at least partially with their binding affinity to an Ago protein, two different siRNA pairs targeting the serine/threonine kinase 1 (PLK1) mRNA were designed and characterized in preliminary experiments[1]. From these preliminary experiments, a pair of siRNAs was chosen whereby one siRNA had a high knockdown efficiency and a strong phenotype (si428) and the second one had lower knockdown efficiency together with a weak phenotype (si845).

The single strands of the siRNAs were synthesized as a 3' Cy5 labeled guide strand and an unlabeled passenger strand (Eurogentec) and the duplex was generated by an annealing reaction. Furthermore, a lysate was needed, where Ago adapted a native conformation and that was free from aggregates, to be able to investigate the binding of the labeled duplex in solution. For the generation of such a lysate, several conditions were tested resulting in one lysate that was suitable for the measurement. This lysate was based on a cytoplasmic fractionation (Dignam et al., 1983a) with cells from a stable and inducible overexpression human embryonic kidney (HEK) T-REx™-293 cell line and which was supplemented with 2 mM $MgCl_2$, 100 mM KCl and 5 % glycerol for stabilization of the respective Ago protein (see Methods chapter 4.2). Taken these findings together, a lysate was generated where the binding of GFP Ago to a Cy5 labeled siRNA could be measured and that was also stable upon several freeze thaw cycles.

[1] See author-contribution page

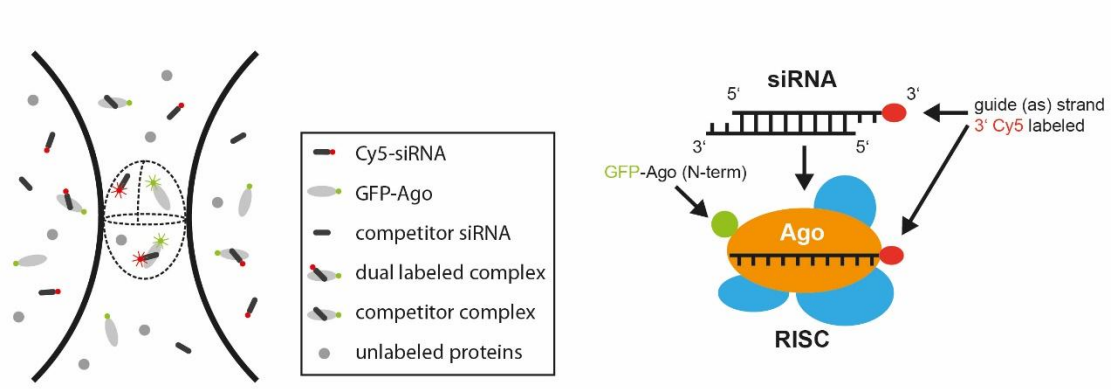


Figure 2.1: FCCS based loading assay. Two laser-lines illuminate a microscopic detection volume and molecules entering or leaving this volume give rise to a fluctuation curve which contains various biophysical data of the involved components (left). For the FCCS assay, Ago was fused to GFP and the siRNA was labeled with Cy5 on its guide 3' end (right).

To further validate if the fluorescent labels of all components have an impact on the functionality, several experiments were performed. As an initial experiment the migration of GFP-Ago2 in a 15-55 % sucrose gradient was investigated. As shown in Höck et al. (2007), Ago2 migrates in three distinct complexes, where each complex differs in Dicer and RISC activities. Therefore, a HEK293 lysate containing overexpressed GFP-Ago2 was loaded onto the gradient and samples from each fraction were analyzed via western blot with the anti-Ago2 11A9 antibody. As expected, the distribution of GFP-Ago2 resembles the wildtype protein, indicating that the GFP-tag does not influence the formation of the different Ago complexes (Figure 2.2 A).

In a second experiment[1], the slicer activity of GFP-Ago2 has been investigated as described in Meister et al. (2004). FLAG/HA (FH)-tagged GFP (Figure 2.2 B lane 2) and GFP-Ago2 (lane 3) were immunoprecipitated with an anti-GFP antibody, FH-Ago2 (lane 5) with an anti-FLAG antibody and all samples were incubated with a radiolabeled target RNA fully complementary to endogenous miR-19b. Although the N-terminus together with the PIWI domain contributes to the slicing activity of Ago2 (Hauptmann et al. 2013) a GFP-tag in this region does not interfere with the cleavage efficiency of the RNA target as the cleavage product is generated in similar amounts compared to FH-Ago2.

Finally the functionality of the labeled siRNAs was verified in a control knockdown experiment in order to exclude, that the Cy5 dye used for labeling of the 3' end of the guide strand of the siRNA interferes with the incorporation into active RISC. Due to the strong regulatory involvement of PLK in the cell cycle, the knockdown was performed only for 24 h to prevent cell-arrest or -death. As shown in Figure 2.2 C, the relative mRNA levels decrease upon siRNA knockdown, where Cy5-si428 has a high knockdown efficiency of ~92% and Cy5-si854 has a lower knockdown efficiency of ~60 % reduction of relative mRNA level, normalized to

[1] See author-contribution page

glyceraldehyde 3-phosphate dehydrogenase (GAPDH) mRNA expression and a control knockdown with an control siRNA. Therefore, the two labeled siRNAs were used to test if knockdown efficiencies of siRNAs correlate at least partially with their binding affinity to Ago proteins.

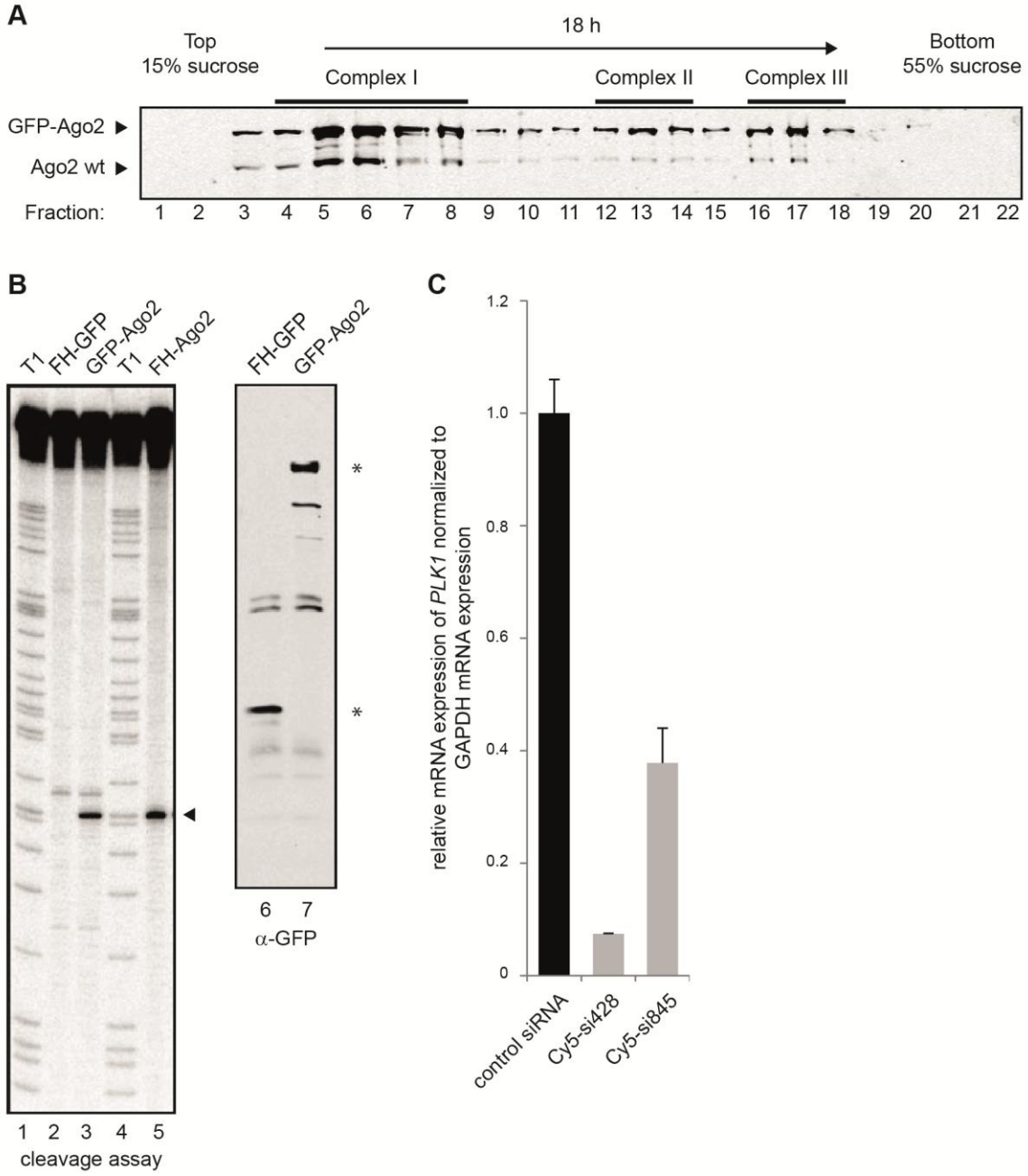


Figure 2.2: Establishment of the FCCS assay (A) Sucrose density gradient for analysis of the migration of GFP-tagged Ago2 into different complexes (complexes I-III, lanes 3-18) confirmed by an Ago2 western blot. **(B)** *In vitro* Ago2 cleavage assay (lanes 2, 3 and 5). GFP-Ago2 was precipitated from a lysate of stable GFP Ago2 expressing HEK293 TRex cells using an GFP-specific antibody and incubated with a fully complementary radiolabeled target RNA to the endogenous miR-19b. As a marker a T1 digestion of the radiolabeled target was loaded (lane 1 and 4). The cleavage product is marked with an arrow (left panel) and FH-GFP (lane 6) and GFP-Ago2 (lane 7) on the loading control western blot (right panel) with an asterisk. **(C)** A knockdown of human *PLK1* was used to validate the functionality of the Cy5 labeled siRNAs using quantitative real-time PCR. Samples were normalized to GAPDH mRNA expression and a control knockdown with a control siRNA.

In the FCCS binding assay, the K_d values of the Cy5 labeled siRNA were determined with a titration of the labeled siRNAs in GFP-Ago2 S100 lysate by serial dilutions. As expected, the siRNA gets incorporated into GFP-Ago2 and the K_d value for Cy5-si428 could be calculated as 126 ± 15 nM (Figure 2.3 A) and for Cy5-si854 as 544 ± 1 nM (Figure 2.3 B), indicating a weaker affinity to GFP-Ago2 of the siRNA with the lower knockdown efficiency. This difference derives from a thermodynamic less stable U/A pair from 5' end of the guide strand from Cy5-si428 compared to a C/G pair of the guide strand from Cy5-si856 which has been shown to be important for efficient Ago loading (Khvorova et al., 2003; Schwarz et al., 2003).

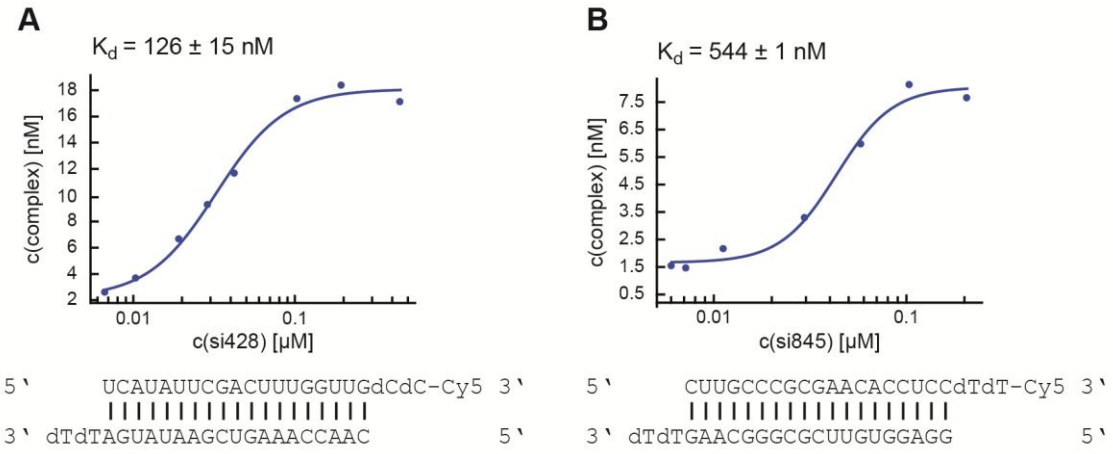


Figure 2.3: Binding of fluorescently labeled siRNAs to Ago2. (A) FCCS titration curves of Cy5-si428 (high knockdown efficiency) and (**B**) Cy5-si845 (weak knockdown efficiency) indicating different affinities of both siRNAs to GFP-Ago2.

Summarized, a lysate condition could be identified where the binding of a labeled siRNAs to GFP tagged Ago2 can be monitored. Due to its higher affinity to Ago2, the Cy5-si428 was used for the following binding studies.

1.2 Loading of Ago1-4

After the characterization of the binding of Cy5-si428 to GFP-Ago2, the FCCS loading assay was applied to the remaining Ago proteins. Therefore stable inducible monoclonal HEK T-REx^{TM}-293 cell lines were established expressing GFP-Ago1, -Ago3 and a polyclonal line for GFP-Ago4. As Ago1, Ago3 and Ago4 were not able to cleave the passenger strand, additionally to the GFP-Ago2 cell line, a stable, polyclonal GFP-Ago2 H807R was generated where the catalytic DDH-motif is mutated to DDR abolishing the slicer activity of Ago2. This mutation was introduced to characterize the influence of the slicer activity to the binding of Cy5-si428 to an Ago protein by directly comparing slicer active and inactive Ago variants. After induction with tetracycline and expression for 24 hours, lysates were generated. The binding of Cy5-si428 to the different GFP-

Ago constructs was measured in the FCCS assay and K_d values from the different Ago proteins were calculated based on the respective titration curves.

First the affinity of Cy5-si428 to GFP-Ago1 was measured resulting in a K_d value of 288 ± 70 nM (See Figure 2.4). For GFP-Ago2 wt a K_d of 217 ± 4 nM and for the catalytic mutant H807R a K_d of 230 ± 12 nM could be obtained showing no difference in the affinity of the labeled Cy5-si428 duplex to Ago2 in a slicer active compared to a slicer inactive Ago protein. In case of Ago3, FCCS measurements resulted in a K_d of 76 ± 14 nM, which was the strongest affinity measured for Cy5-si428. The measurement of GFP-Ago4 resulted in a K_d of 148 ± 31 nM. Taken together, these results show that the FCCS assay is applicable to all GFP-labeled Ago proteins and binding of each protein could be observed. Furthermore, all Ago proteins seem to have similar affinities to Cy5-si428 and variabilities of the measured K_d values could derive from the quality of the lysates (see Discussion chapter 1.3).

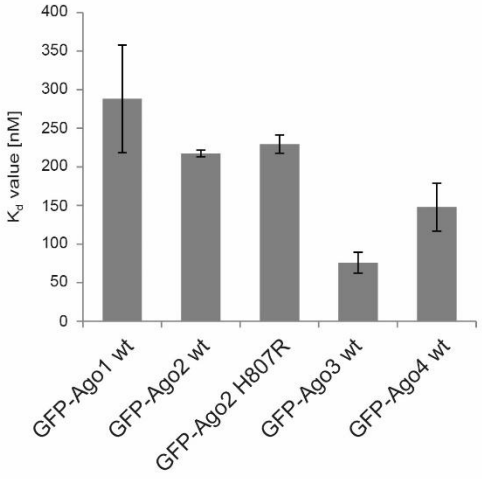


Figure 2.4: Overview of the K_d values from GFP-tagged Ago proteins. Columns represent K_d values of different Ago proteins. Error bars were obtained from replicates of several measurements of each titration curves.

1.3 Different siRNAs bind to Ago with different affinities

To further validate the hypothesis whether differences in the affinities of siRNAs might be one determinant for an efficient siRNA knockdown, a competition assay was designed, to get access to the affinities of any desired siRNA. For this, increasing amounts of unlabeled siRNAs were mixed with a constant amount of Cy5-si428 and added to the sample to allow competition for GFP-Ago2 binding. A loss of cross-correlation between GFP-Ago2 and Cy5-si428 can be measured depending on the concentration of the added unlabeled siRNA. Using this approach, the K_i values, which are the K_d values for the competitor siRNA binding reflecting their respective affinity to GFP-Ago2, of eight different siRNAs (Figure 2.5 A) were determined. The siRNAs were designed against members involved in the human nuclear import pathway (Importin β

(Impβ), Importin 11 (Imp11), Ran-specific GTPase-activating protein (RanBP1) and Transportin-1 (Tnpo1)) and already characterized concerning their knockdown efficiencies (Schraivogel et al., 2015). The repetition of the knockdown experiment, using quantitative real-time PCR (Figure 2.5 B), could confirm the efficiencies described in this publication.

A comparison of the affinities derived from the FCCS assay with their respective *in vivo* knockdown efficiencies showed at least some correlation between both experiments. Impβ si1 and si2, Imp 11 si1, Tnpo1 si1and RanBP1 si1 and 2 altogether led to a strong reduction of their respective target mRNA to less than 20% (Figure 2.5 B). Consistent with that, their respective K_i values showed a relatively high affinity to GFP-Ago with ~30nM (Figure 2.5 C). Tnpo1 si2, which comprised the most inefficient knockdown efficiency with a residual mRNA level of 73% (Figure 2.5 B), also had the weakest binding affinity to GFP-Ago2 in the FCCS assay with a K_i value of 76 nM (Figure 2.5 C), which was more than two times higher than the K_i values obtained for an efficient siRNA. Regarding only the Tnpo1 siRNA pair, which are designed against the same target mRNA, the difference of the knockdown efficiency might be influenced by the affinity of the siRNA to GFP-Ago2. However regarding the Imp11 si1 and 2 pair, Imp si2 has a lower knockdown efficiency of about 34 % residual mRNA level but a similar affinity to GFP-Ago2 with a Ki value that is similar to Imp11 si1 which has a knockdown efficiency of ~15% residual mRNA level.

Taken these findings together, a competition between Cy5-si428 and any unlabeled siRNA for GFP-Ago2 binding can be measured with the FCCS assay, resulting in the respective K_i values of each single unlabeled siRNA. For two different siRNAs targeting the Tnpo1 mRNA, a correlation between the knockdown efficiency and the affinity to GFP-Ago2 was measured, however for statistical validation of the hypothesis that the knockdown efficiency of a certain siRNA correlates, at least to some extent, with its affinity to Ago2, a dataset of several more siRNAs would be required.

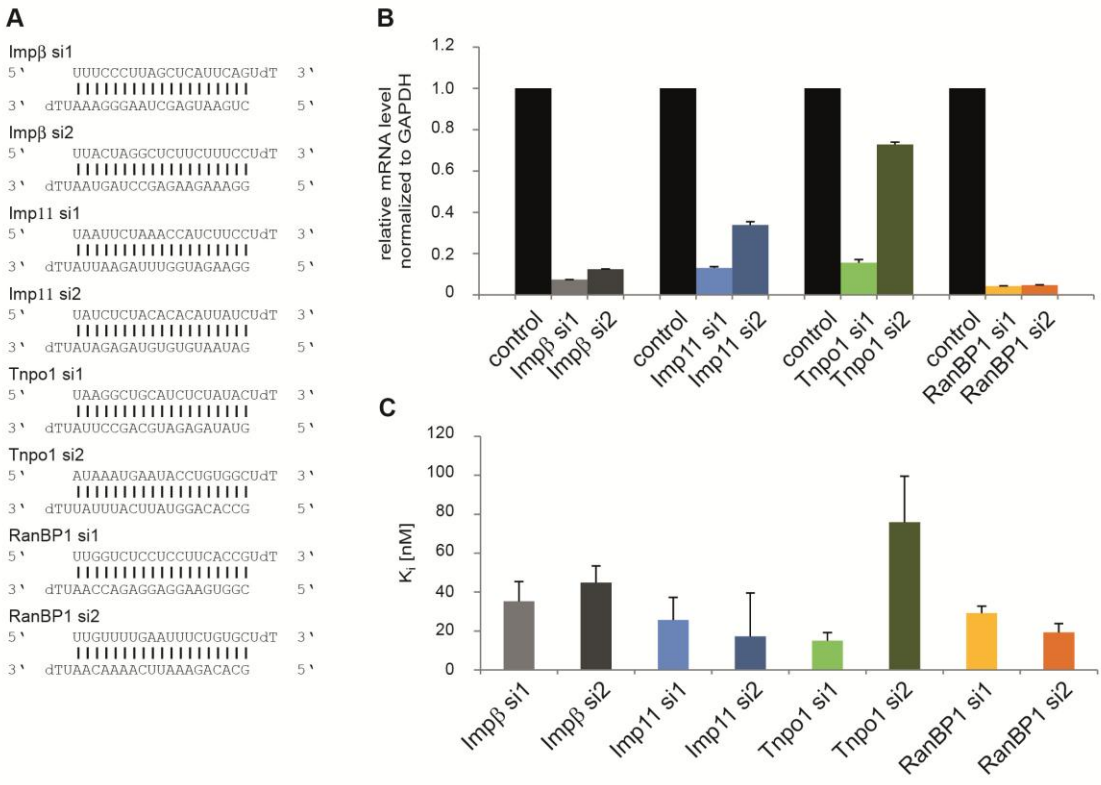


Figure 2.5: siRNA competition screen. (A) Sequences of different siRNAs designed against members involved in the human nuclear import pathway. **(B)** Quantitative real-time PCR analysis of knockdown efficiencies from different siRNAs designed against members involved in the human nuclear import pathway, normalized to GAPDH mRNA levels. **(C)** K_i values from different unlabeled siRNAs derived from an FCCS competition screen with Cy5-si428 and GFP-Ago2.

1.4 Influence of Ago phosphorylation on siRNA loading

Many proteins in the cell are post-translationally regulated via dynamic phosphorylation of serines, threonines and tyrosines. Several phosphorylation (phospho)-sites have been reported for endogenous and overexpressed Ago2 (Rüdel et al., 2011) and (Sharma et al., 2014). As the FCCS assay has been shown to be a powerful tool for the investigation of the binding between GFP-tagged Ago variants and the Cy5-labeled si428, the influence of two potential phospho-sites within the sequence of Ago2 were measured. Therefore, stable cell lines expressing GFP-tagged phospho-mimics were generated and lysates were titrated with Cy5-si428.

In case of Y529 a phosphorylation affects the binding of endogenous miRNAs (Rüdel et al., 2011). Y529 is located in the Mid domain of Ago2 and is required for binding of the 5' end of a miRNA via its sidechain aromatic ring and hydroxyl group. The aromatic ring of Y529 has a parallel orientation to the first base of the miRNA with a distance of ~3.7 Å which allows π-stacking interactions and stabilizes the "kinked out" conformation of the first base (Figure 2.6 A). The Hydroxyl group further stabilizes this conformation via a hydrogen bond with the backbone phosphate of the first nucleotide (Figure 2.6 A). Mutations of Y529 to alanine, glutamate,

phenylalanine and glutamine were described in Rüdel et al. (2011). These findings were used as a proof of principle of the biophysical data generated in the FCCS assay.

For this purpose, first the expression level of each single mutant in the S100 lysate was controlled using western blots against GFP and tubulin (Figure 2.6 B). All mutants and the wildtype Ago2 were equally expressed indicating that the mutations have no influence on their respective expression level.

Then the mutants were measured in the FCCS assay starting with a S100 lysate containing overexpressed GFP-Ago2 Y529E. Figure 2.6 B shows the titration curve where the concentration of the Cy5-si428 has been plotted in logarithmic scale. As expected, no binding at any Cy5-si428 concentration and therefore no K_d value could be measured. The negatively charged side chain of glutamate repulses the 5' phosphate of the siRNA and efficiently prevents binding into the binding pocket of Ago2. Thus, a phosphorylation at Y529 might have a post-transcriptional regulatory effect on Ago proteins. Therefore the observation of the FCCS assay is consistent with the findings described in Rüdel et al. (2011).

If Y529[1] was replaced by an alanine, the K_d value increases from 178 ± 1 nM measured for the wt GFP-Ago2 to 1.56 ± 0.36 µM (Figure 2.6 D). Y529 stabilizes the kinked conformation of the first nucleotide of the miRNA, where the base stacking with the second nucleotide is interrupted, via a hydrogen bond with the 5' phosphate and π-stacking with the base. Alanine is not able to stabilize the 5' end of the miRNA, which might explain the shift of the K_d values.

A mutation of Y529 to glutamine introduces a bulky polar side chain into the 5' binding pocket of Ago2. Consistent with that, the K_d is affected even stronger than in the Y529A mutant and is shifted to 2.61 ± 0.38 µM. The polar side chain of glutamine sterically interferes with the 5' end of the miRNA and is also not able to stabilize the kinked conformation. But in contrast to the Y529E mutant, where a strong electrostatic repulsion of the negatively charged side chain and the 5' phosphate completely abolishes the binding, Y529Q is still able to bind the labelled siRNA, although with a very low affinity.

A mutation of Y529 to phenylalanine has only a small influence on the affinity. The K_d could be determined as 209 ± 7 nM and is almost identical to the K_d of the wildtype Ago2 (178 ± 1 nM). As the 5' phosphate of the miRNA is not only coordinated by Y529 but also via hydrogen bonds to K533, Q545, K566, K570 and R812, Y529 might contribute more via π-stacking to the binding then via its hydrogen bond. Thus Y529F, which lacks the contribution of the hydrogen bond, is still able to stabilize the kinked out conformation via π-stacking of its aromatic ring to the first base of the miRNA.

Taken these findings together, the FCCS assay was able to reproduce findings on the potential phospho site Y529 (Rüdel et al., 2011) and kinetic data in terms of the respective K_d values could

[1] See author-contribution page

be added within this study. Furthermore, the FCCS assay has proven to be suitable for the measurement of any potential phospho site within the sequence of an Ago protein.

Thus another less characterized phospho site, which was additionally identified in the mass spectrometric analysis from Rüdel et al. (2011) was subjected to the FCCS assay to validate whether it could be a candidate for a possible phosphorylation site on Ago2 with a functional relevance on siRNA binding. S798 is located in a flexible loop region within the PIWI domain and contacts the RNA via a hydrogen bond to the phosphate group of nucleotide 5 within the miRNA sequence (Figure 2.6 A). To further validate the influence of S798 on the binding to dsRNA, the GFP-Ago2 mutants S798A and the phospho-mimic S798E were generated.

For the S798A mutant a K_d value of 155 ± 27 nM could be measured indicating no effect on the binding of the labelled siRNA. The removal of one single hydrogen bond to the backbone phosphate of nucleotide 5 is not sufficient to reduce the affinity and is completely compensated by residual interactions of Ago2 to the siRNA. The phospho-mimic S798E with a K_d value of 206 ± 12 nM seems to have a small decrease of the affinity. However, this could also be explained by variations of the FCCS assay. The flexibility of the loop where S798 is located might also be responsible for the compensation of the negative repulsion from the phosphate backbone and the glutamate side chain. A phosphorylation at this residue alone thus has no post translational regulatory influence on the loading of Ago2.

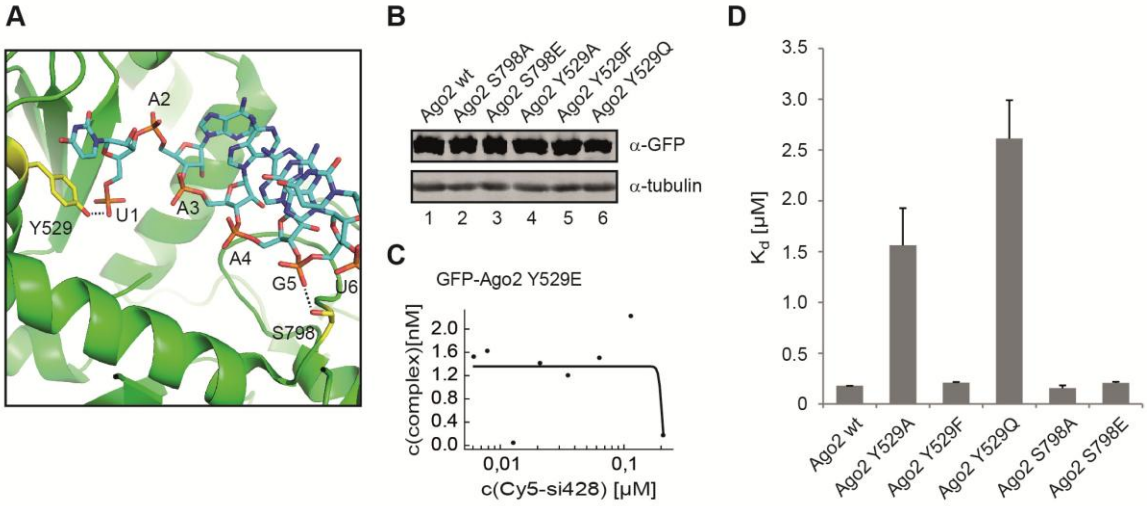


Figure 2.6: Ago2 phospho-mutants. (A) Overview of the 5' miRNA-binding pocket of Ago2 in complex with miR-20a (PDB 4F3T). Sidechains of Y529 and S798 which have been mutated for the binding assay are shown as yellow sticks. **(B)** Loading and expression control of lysates from polyclonal phospho-mutants via western blot against GFP and tubulin. **(C)** Titration curve of the binding deficient GFP-Ago2 mutant Y529E. **(D)** K_d values of different phospho-mutants tested in the assay.

2. The fly miRNA biogenesis factor Loqs

In the conserved miRNA biogenesis pathway, endogenous miRNAs have to pass several processing steps before they were loaded to their terminal Ago effector proteins. After passenger strand removal the miRNA single strand is able to perform translational repression of its target mRNA. The Dicer processing step, where the miRNA precursor is cleaved and the mature miRNA is released, directly precedes the Ago loading step and is assisted by a dsRBP partner. In flies, where two distinct Dicers exist, this step is carried out by dmDcr-1 and the Loqs-dsRBD3. Thereby the third dsRBD of Loqs has an identical fold from a classical dsRBD but functions as a protein interaction domain (Förstemann et al., 2005) instead of binding to RNA. To shed light on the molecular and functional details underlying the dmDcr-1-Loqs interaction, we structurally and functionally characterized important details of the dmDcr-1-Loqs interaction based on the crystal structure of Loqs-dsRBD3.

2.1 Expression of Loqs-dsRBD3

The boundaries of the Loqs-dsRBD3 crystallization-construct were determined according to a secondary structure prediction (PredictProtein) and a sequence coding for the isolated dsRBD3 (residues 392-463) was cloned into expression vectors for initial expression- and solubility-screening. The optimal expression of a soluble Loqs-dsRBD3 was achieved in *E. coli* BL21 (DE3) with the expression vector pET32a by overnight expression at 18°C. Figure 2.7 shows the expression of the Loqs-dsRBD3 at different temperatures and time points.

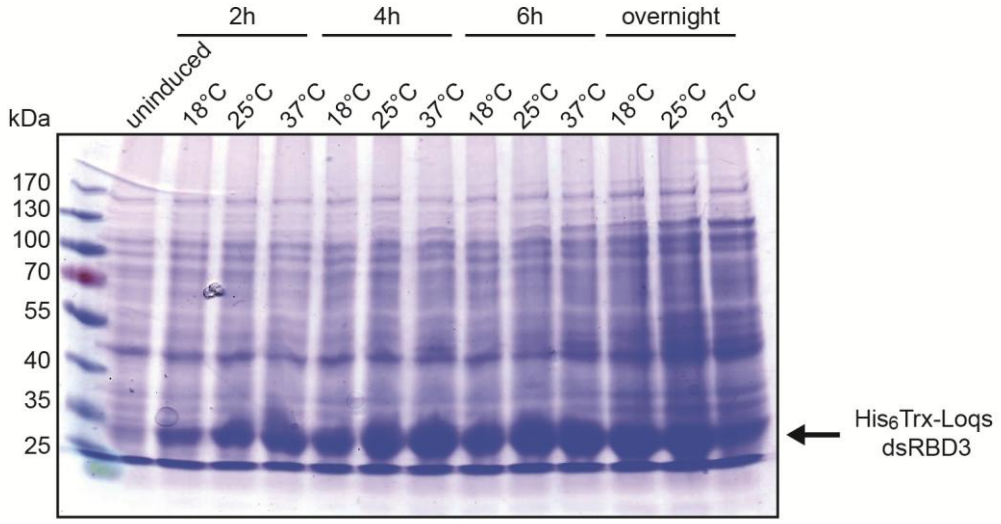


Figure 2.7: Expression of His_6Trx-Loqs-dsRBD3. After induction with IPTG at OD_{600}=0.6, cells were grown at 18, 25 and 37°C and samples were taken after 2h, 4h, 6h and overnight. Proteins were separated on a 10 % SDS gel and stained with Coomassie.

2.2 Purification of Loqs-dsRBD3

E. coli BL21 (DE3) were transfected with the pET32 vector containing an N-terminal His_6-Thioredoxin (Trx) tag for affinity chromatography purification and an overnight culture was inoculated form these transfected cells. From that culture, lysogeny broth (LB)-medium was inoculated and grown overnight at 18°C after induction with IPTG (OD_{600}=0.6). Cells were harvested and proteins were extracted using sonication. A lysate was obtained, containing large amounts of the overexpressed protein, which was affinity purified by immobilized metal ion affinity chromatography (IMAC) (Figure 2.8, A). Elution fractions were pooled and subjected to a TEV digestion while dialyzing against HisA buffer. Due to the inefficient cleavage of the TEV protease, which probably resulted from a steric hindrance between two folded protein regions, a linker of two additional glycines was introduced at the N-terminus of the construct and the cleavage activity was improved to ~60% of the total protein (Figure 2.8, B).

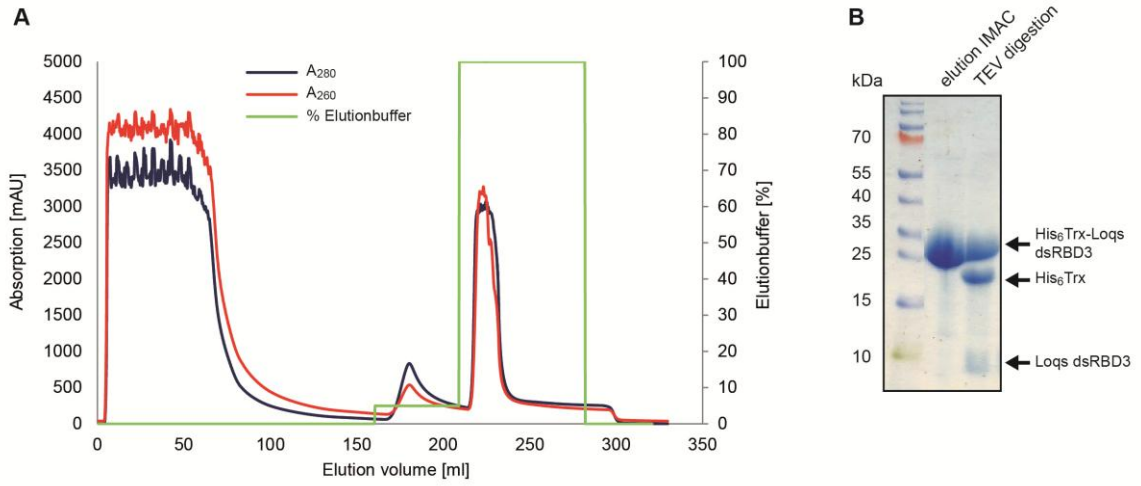


Figure 2.8: First IMAC of His_6Trx-Loqs-dsRBD3. (A) Chromatogram of the IMAC of His_6Trx-Loqs-dsRBD3. **(B)** 15 % SDS gel of the pooled fractions from the elution peak (~225 ml) of His_6Trx-Loqs-dsRBD3 and protein composition after TEV digestion of the respective pooled fractions.

The uncut protein, the Trx-His_6 tag and the TEV protease were removed with a second IMAC (Figure 2.9, A) and the flow through was concentrated by ammonium-sulfate precipitation (50% w/v) and centrifugation. The flow through of the second IMAC contained already high amounts of the isolated Loqs-dsRBD3 (Figure 2.9, B). However a minor proportion of the uncut protein and the isolated His_6Trx-Tag co-eluted with the isolated dsRBD3. Therefore size exclusion chromatography (SEC) was applied after resuspension of the precipitated protein pellet.

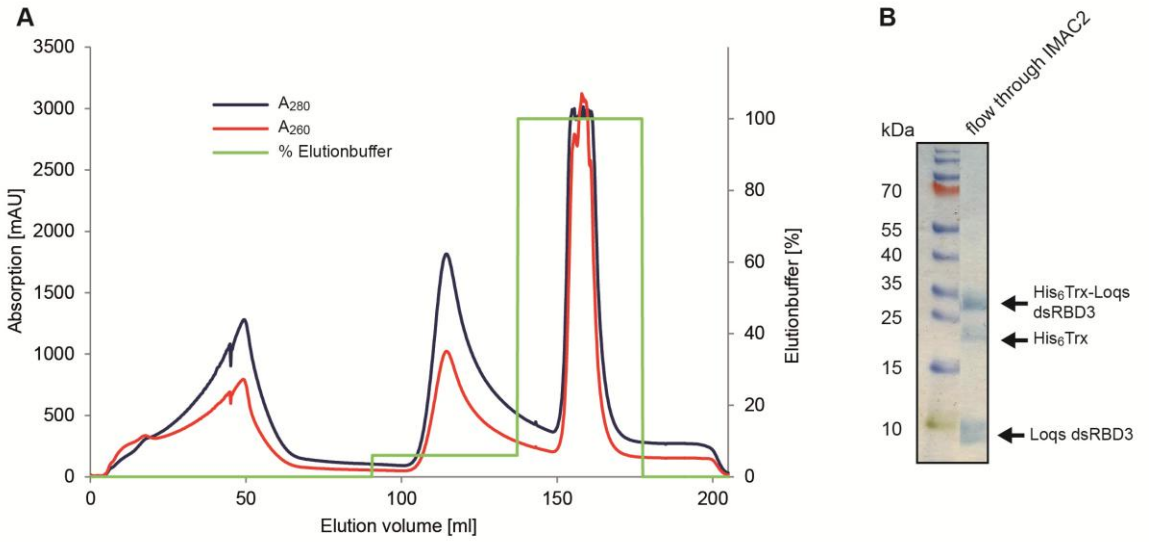


Figure 2.9: Second IMAC of His_6Trx-Loqs-dsRBD3. (A) Chromatogram of the second IMAC of TEV cleaved His_6Trx-Loqs-dsRBD3. **(B)** 15 % SDS gel of the pooled fractions from the flow through peak (~5-60 ml).

After the SEC run, the Loqs-dsRBD3 with a molecular weight of ~8 kDa eluted at ~188 ml, which was unexpectedly early (See chapter 2.8 for further information) for a protein of the respective size (Figure 2.10, A). The shoulder from the elution peak was referred to the His_6Trx-Tag with a molecular weight of ~18 kDa, which eluted at a volume of ~174 ml. The peak fractions were pooled and separated from the shoulder and the purity of the respective fractions was confirmed with a 15 % SDS gel that was stained with Coomassie (Figure 2.10, B). After the first SEC, still a very little residual proportion of the His_6Trx-Tag was present in the elution fractions. Therefore fractions were concentrated and subjected to another SEC.

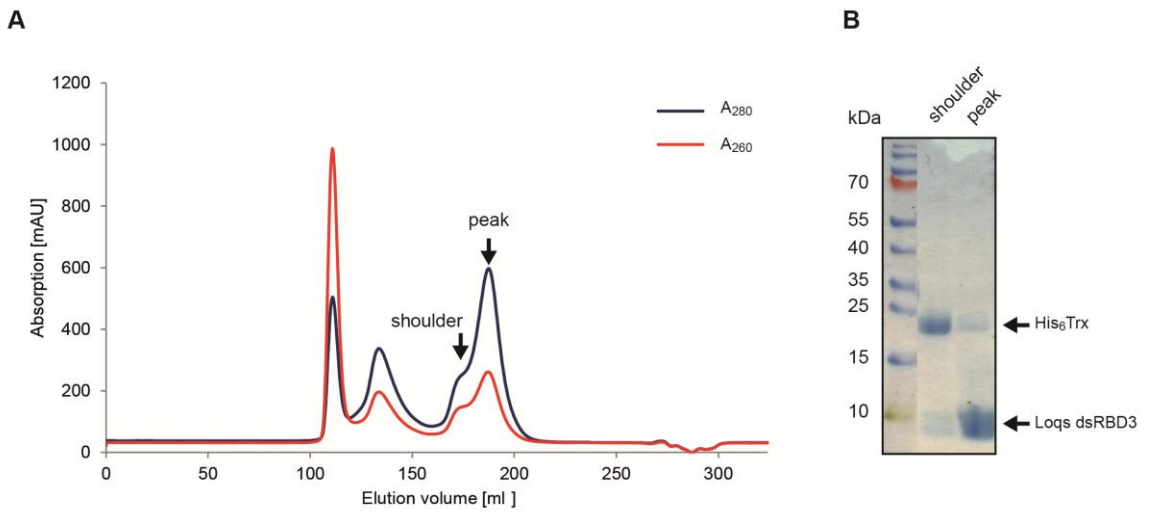


Figure 2.10: First SEC of Loqs-dsRBD3 (A) Chromatogram of the first SEC Loqs-dsRBD3. **(B)** 15 % SDS gel of the pooled fractions from the elution peak (~188 ml) and the shoulder (~174ml).

After the second SEC, highly purified Loqs-dsRBD3 was obtained which eluted in a single homogeneous peak at ~188 ml (Figure 2.11, A). The SDS gel from the pooled fractions showed a clear band for the isolated Loqs-dsRBD3 with no further contaminations (Figure 2.11, B). Therefore, the highly purified Loqs-dsRBD3 was used for crystallization experiments.

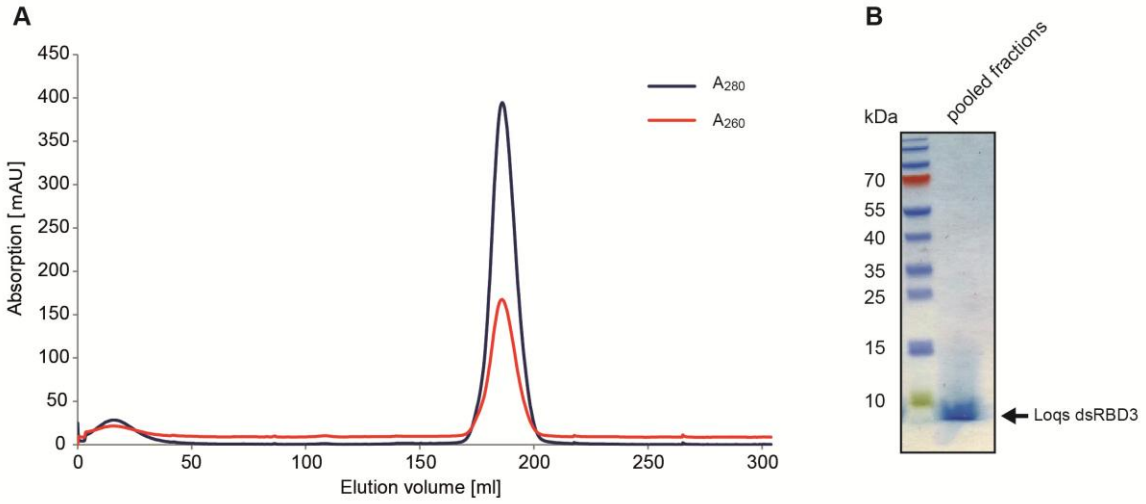


Figure 2.11: Second SEC of Loqs-dsRBD3 (A) Chromatogram of the second SEC Loqs-dsRBD3. **(B)** 15 % SDS gel of the pooled fractions from the elution peak (~188 ml).

2.3 Crystallization of Loqs-dsRBD3

After the final SEC, the buffer of Loqs-dsRBD3 was exchanged to crystallization buffer (20 mM Tris pH 7.5, 2 mM DTT) and the protein concentration was adjusted to ~15 mg/ml. Then screening was performed with commercially available crystal screens (See Table 4.11) using the "sitting drop" method (see Methods, chapter 7.1). However no condition for initial crystallization could be identified from all tested screens.

Therefore another strategy was applied where the Loqs-dsRBD3 was chemically modified (Walter et al., 2006). As lysines can have a negative steric effect upon crystal contact formation due to a long positively charged sidechain, all free amino groups (Lysine side chains and the N-terminal amino group) were di-methylated within the sequence of Loqs-dsRBD3. The efficiency of the reaction was validated with a mass spectrometry analysis showing all four amino groups (3 lysines and the N-terminus) being methylated. This is reflected by a shift of the full length Loqs-dsRBD3 in the MS spectra, from a measured molecular mass of 8088.9 Da for the wildtype protein (Figure 2.12, A) and for the fully methylated protein to 8200.7 Da (Figure 2.12, B). To remove the chemical components from the protein methylation, the fully methylated construct was subjected in to another SEC, where a shift of the elution volume from 188 ml for the unmethylated protein to 181 ml for the fully methylated protein occurred (see Figure 2.12 C). The purity of the methylated Loqs-dsRBD3 was controlled by loading different amounts of the protein

onto a 10% Schägger gel and subsequent Coomassie staining (Figure 2.12 D). No additional bands were visible except of a faint band at ~16kDa which was identified by another MS-analysis (data not shown) as the dsRBD3 construct and might represent a SDS-stable dimeric form of the protein.

In a last chromatographic step, the buffer was exchanged to crystallization buffer and immediately used for crystallization screening with commonly available crystal screens (Table 4.11). From crystal screening of the methylated protein, one single condition for crystallization could be identified consisting of 10 % poly ethylene glycol (PEG) 1000 and 10% PEG 8000. After refinement screening using the hanging drop method, the optimal mixture for crystallization of both precipitants could be identified as 2 % PEG 1000 and 12 % PEG 8000 supplemented with 5 mM DTT which was crucial for the formation of the crystals. 2µL from the reservoir were mixed with 2 µL of a ~15 mg/mL protein solution and crystallized at 20°C. After 4-7 days, long and rod-shaped crystals grew within the drop (Figure 2.12, E). In most of the cases, crystals grew from a solid base with a cavity ranging from one side of the rectangle to the middle of the crystal. For further structural analysis, single crystals or single parts of the bases of one crystal were fished and briefly soaked in 20 % glycerol, 2% PEG1000 and 12% PEG8000 for protection against cryo-damage during freezing in liquid nitrogen. After the incubation with cryo buffer, the crystals were frozen and stored in liquid nitrogen.

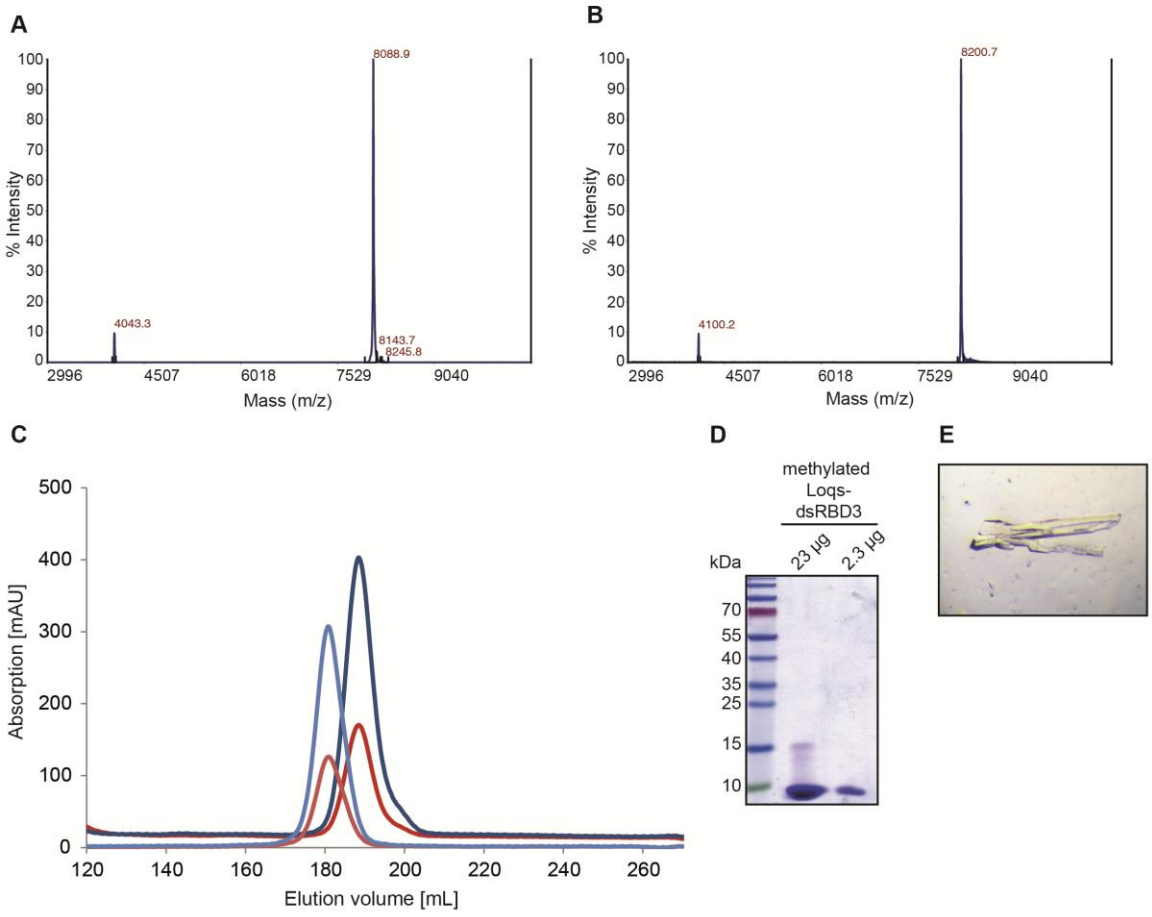


Figure 2.12: Crystallization of methylated Loqs-dsRBD3. (A) MS spectra of unmethylated Loqs dsRBD3 **(B)** MS spectra of methylated Loqs dsRBD3 **(C)** Size exclusion chromatography of the unmethylated (dark blue (A280nm) and red (A260nm)) and the methylated (light blue (A280nm) and light red (A260nm)) dsRBD3 of Loqs. The protein elution shifts to a lower elution volume upon addition of methyl groups to the four amino groups present in the construct. **(D)** 10 % Schägger gel showing different amounts of the highly purified methylated protein with a molecular weight of 8.2 kDa and a SDS stable dimeric form of the protein at ~16 kDa. **(E)** Picture of a crystal obtained after refinement screening.

2.4 Data collection and structure determination

The data collection was performed at the Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung m.b.H. (BESSY) on the beamline BL14.2 where native and selenomethionine (SeMet) substituted protein crystals (Methods chapter 7.2) were measured. SeMet incorporated Loqs-dsRBD3 was purified and crystallized in a similar manner as shown for the Loqs-dsRBD3.

For the measurements, high energy synchrotron radiation was used, which is a prerequisite for X-ray crystallography. This radiation, like every electromagnetic wave, consists of an amplitude, that can be measured in form of the intensity of each spot on the detector and a certain phase, that is systematically lost upon diffraction at the electrons from atoms within the crystal. To overcome the phase problem, multi-wavelength anomalous diffraction (MAD) was performed. First, a fluorescence scan with the SeMet crystals was performed, for the detection of the optimal

wavelength for the three MAD measurements (peak, inflection and high remote, data not shown). Crystals from the SeMet protein diffracted up to 2.9-2.7 Å resolution and for the native protein up to 2.65 Å. The measured datasets were processed using XDS (Kabsch, 2010). Statistics of the MAD and the native datasets are given in Table 2.1.

Table 2.1: Data collection statistics of Loqs-dsRBD3 crystals

	peak	inflection	high remote	native
Wavelength (Å)	0.9798	0.9800	0.9082	0.9184
Resolution range (Å)	50-2.9 (3.08-2.9)	50–2.7 (2.86-2.7)	50–2.9 (3.08-2.9)	50–2.65 (2.81-2.65)
Space group		$P\ 2_1\ 2_1\ 2_1$		$P\ 2_1\ 2_1\ 2_1$
Unit cell, a b c		65.8 112.7 114.4		66.0 113.0 (114.1)
Total reflections	285192 (45395)	347905 (51885)	143778 (22934)	188349 (30280)
Unique reflections	36232 (5790)	44625 (6932)	36379 (5825)	25522 (4047)
Multiplicity	7.9 (7.8)	7.8 (7.5)	4 (3.9)	7.4 (7.5)
Completeness (%)	99.7 (99.1)	99.0 (95.1)	99.7 (99.4)	99.9 (99.9)
Mean I/sigma(I)	23.0 (5.4)	27.7 (6.2)	21.2 (5.1)	33.8 (6.0)
Wilson B-factor [$Å^2$]	47.4	55.3	49.8	54.3
R_{sym-I} [%]	8.9 (48)	6.5 (45)	5.6 (30.3)	4.6 (35.2)

Values as reported by *CORRECT*. Values in parentheses refer to the highest resolution shells.

Based on the Matthews coefficient, which allows an estimation of Loqs-dsRBD3 monomers that are present in the asymmetric unit by comparing the volume of the asymmetric cell with the molecular weight of the Loqs-dsRBD3 and the number of symmetry operators in the space group $P2_12_12_1$, a number of ˜8 monomers and a relatively high solvent content was obtained for the Loqs-dsRBD3 asymmetric unit cell. The high number of molecules within the asymmetric unit cell gave a reason why former attempts using molecular replacement, which is based on the correlation of a previously solved homologous protein structure with the data, obtained for the unknown structure, resulting in a solution for the phase problem, failed to calculate such a solution. However experimental phasing using AutoSHARP (Vonrhein et al., 2007) with the MAD data sets resulted in a total number of 9 selenium sites which was sufficient for the solution of the phase problem.

Six of these sites could be assigned to the C-terminal methionine from each of the monomers resolved in the crystal. The remaining three positions were identified as alternative conformations from the C-termini showing some flexibility at this region. Starting from to the position of the respective SeMet in the N-terminal helix, model building was manually performed using COOT (Emsley et al., 2010). The refinement of the model was performed with the program

PHENIX.*refine* (Afonine et al., 2012) thereby using non-crystallographic symmetry (NCS) restraints and simulated annealing.

Starting from the SeMet positions at the C-terminal helix of the Loqs-dsRBD3, a structural model of the complete asymmetric unit cell (Figure 2.13, A) was built into the electron density map obtained after experimental phasing. Thereby, half of the monomers in the crystal have undergone a domain swap where the β-strand β3 and the C-terminal helix α2 were exchanged between two monomers connecting both macromolecules by a physical bridge (Figure 2.13, B). Together with two "unswapped" monomers, this tetramer formed the building block for the crystal packing and two of these tetramers put up the asymmetric unit cell. It is likely that this domain swap was only required for crystallization as a compensation for lacking crystal contacts between each single monomer of Loqs-dsRBD3 and therefore can be seen as a crystal artifact which probably has no relevance for the function of Loqs *in vivo*. Consistent with that is also the fact, that during purification of the Loqs-dsRBD3, no species of higher molecular weight were observed, meaning that the tetrameric assembly was formed only upon crystallization.

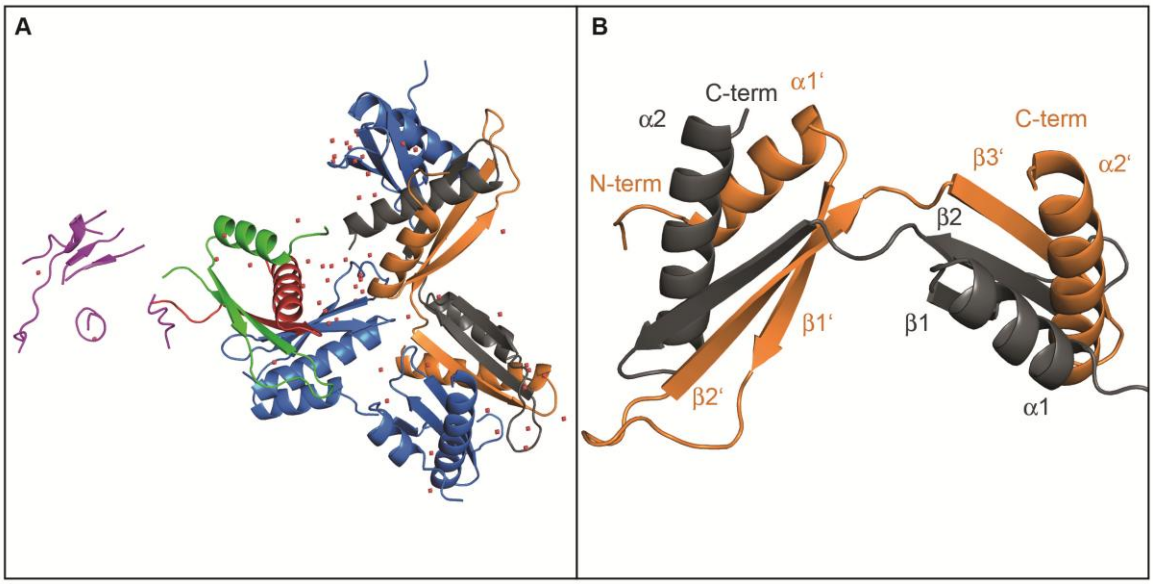


Figure 2.13: Asymmetric unit cell of Loqs-dsRBD3. (**A**) Beside the native monomers (blue) also monomers with swapped secondary structure elements (grey and orange) exist. Another dimer where the domain swap took place is only partly resolved (red and green) and the second half which is protruding in a large solvent channel could only be modeled as a poly alanine trace (magenta). (**B**) Detailed view of the domain swap between two monomers.

One dimer of the building block protrudes into a large solvent channel and does not form any contact to other parts of the crystal. Therefore the electron density in this region is only weakly defined and this is reflected by the high B-factor averaged over all atomic positions within the structural model (Table 2.2). In these parts, the electron density was only partially defined for the main chain and completely missing for the side atoms. Therefore, the main chain was built as poly

alanine trace, as alanine possesses the smallest sidechain of all amino acids and therefore is commonly used, when electron density is missing for side chains (Table 2.2).

Table 2.2: Model statistics and structure refinement

Parameter	Loqs-dsRBD3
R_{work}/R_{free}	20.6/25.1
Modelled polypeptides	
Monomers	
Chain A, C/B	391-462/392-463
"Swapped" monomers	
Chain D/E/F/G	391-463/389-463/429-463/390-433
Fragments	
Chain H	48 amino acids (Ala)
No. of non-hydrogen atoms	3699
macromolecules	3626
water/ligands	67/6
RMS(bonds) [Å]	0.009
RMS(angles) [°]	1.20
Ramachandran favored (%)	97.2
Ramachandran outliers (%)	0
Clashscore	3.62
Average B-factor [$Å^2$]	65.40
macromolecules	65.60
ligands/solvent	79.10/ 52.80
No. of TLS-groups	25

2.5 The crystal structure of Loqs-dsRBD3

Within the asymmetric unit there exist two states for the Loqs-dsRBD3 monomers. One half of the monomers swap secondary structure elements and were physically linked to each other and the other half that do not contain this domain swap and therefore were used for the following structure description.

The Loqs-dsRBD3 has a classical α-β-β-β-α fold, which is a common feature among dsRBDs (Figure 2.14 A). In a superposition of the Loqs-dsRBD3 with other solved dsRBD containing structures (Staufen1, Dicer and RNase III), the atomic positions of the respective amino acids were almost identical with root-mean-square deviation (rsmd) values < 2 Å (data not shown) identified on the DALI server (Holm and Rosenstrom, 2010).

Classical dsRBDs contact dsRNA via three different regions, which are located at one side of the dsRBD, representing the protein part of the RNA interface. Region 1 and 3 are made up by the N- and C-terminal helices and region 2 by the loop L2 between β-sheets 1 and 2 (Figure 2.14 A, yellow). While the residues in region 1 are conserved, there are major differences among the other

two regions. In region 2, the GPxHxx motif is missing and in region 3 the KKxAK motif is mutated to an AAxAQ motif, which lacks the positive charges for the interaction with the major grove of the dsRNA backbone.

The L2 loop, which represents RNA binding region 2 is not only lacking the motif for RNA binding but is also remarkably oriented away from the RNA interface in all monomers of the crystal. Additionally, the length of L2 is shorter compared to RNA binding dsRBDs, which makes it impossible for the Loqs-dsRBD3 to contact and anchor the protein to the minor groove of the dsRNA that is specifically bound by a histidine residue in dsRBD2 of TRBP (Yang et al., 2010). Figure 2.14 B shows a superposition of TRBP dsRBD2 (yellow), which was crystallized together with a dsRNA substrate (Yamashita et al., 2011), Staufen dsRBD5 (blue), which has no affinity to dsRNA and acts, similar to the Loqs-dsRBD3, as a protein interaction domain (Gleghorn et al., 2013) and Loqs-dsRBD3 (green). Consistently, the L2 loop from Staufen1 dsRBD5 like L2 from Loqs-dsRBD3 is shorter and points away from the dsRNA interface, however in a more intermediate orientation (black arrow). Therefore, the orientation and length of the L2 loop might be crucial for either a protein-protein or Protein-dsRNA interacting dsRBD in general. Consistently, a reconstitution of the KKxAK motif in region 3 and the GPxH motif in region 2 is not sufficient to restore RNA-binding activity in gelshift assays, as the shorter L2 loop might prevent the histidine side chain to reach into the dsRNA minor groove (data not shown). In case of TRBP dsRBD2, L2 has exactly the right length, orientation and the GPxH motif and therefore is able to contact the minor groove of the dsRNA to anchor the dsRNA to the protein.

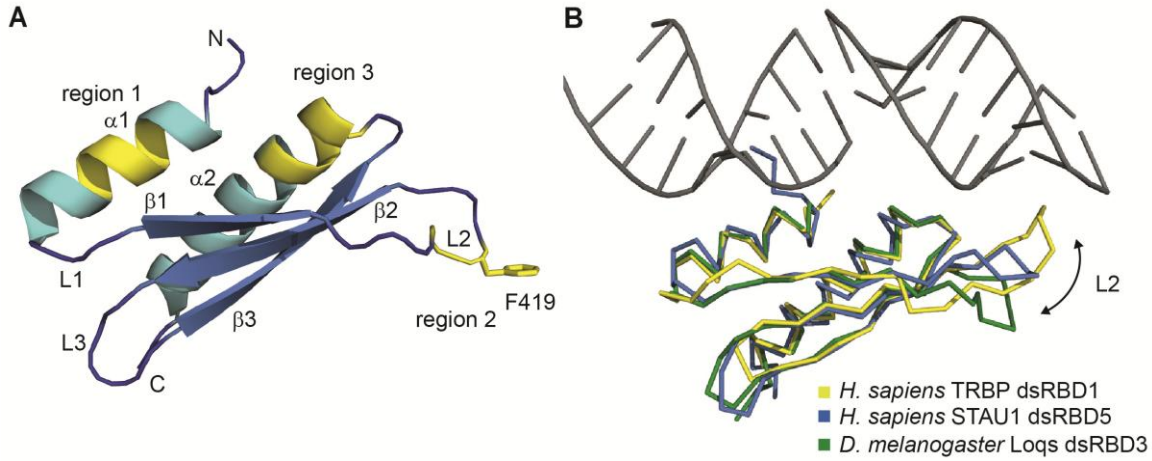


Figure 2.14: Structure of Loqs-dsRBD3. (A) Monomer of the Loqs-dsRBD3. Regions which are responsible for RNA binding in functional dsRBDs were marked in yellow. The side chain of the solvent-exposed phenylalanine (F419) within L2 is shown as sticks. **(B)** Superposition of the structures of TRBP dsRBD2 bound to RNA (PDB=3adl, yellow), RBD5 of Staufen1 (PDB=4dkk, blue), and Loqs-dsRBD3 (green). An arrow highlights different orientations of the L2 loop pointing towards (TRBP) or away (Loqs) from a dsRNA substrate.

2.6 *In vitro* characterization of the Loqs-dmDcr-1 interaction interface

For validation of Loqs-dmDcr1 interaction, glutahione-S-transferase (GST)-pulldown assay were performed. Thereby GST-tagged wt and mutants of Loqs were used as bait proteins and were incubated with a lysate of SF21 cells, containing overexpressed dmDcr-1. Figure 2.15 gives an overview of the GST tagged constructs used for the pulldown assays.

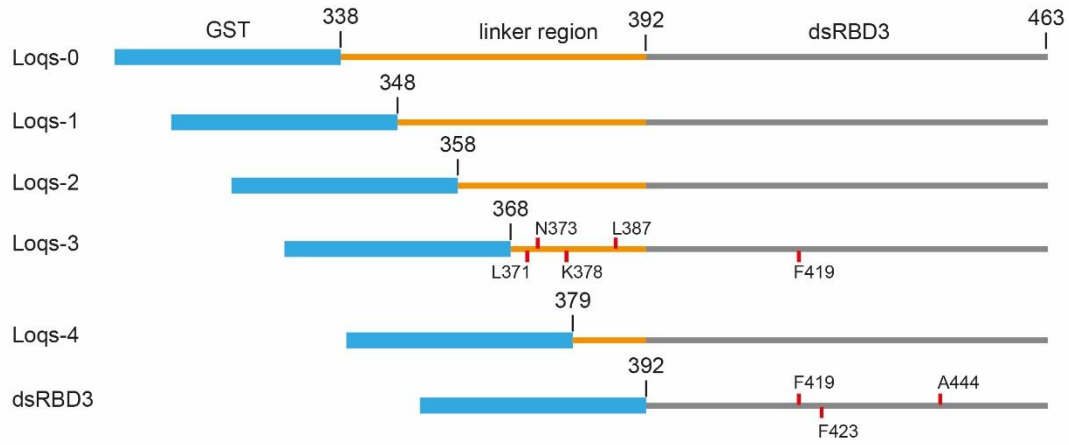


Figure 2.15: Schematic overview of GST tagged Loqs constructs. Positions, where mutations were introduced, are indicated with red lines.

Based on the structure of Loqs-dsRBD3, the most striking position for an amino acid is F419, which is located in L2 and is directly exposed to the solvent in all native species that present in the crystal (Figure 2.14 A). Thus, this residue was subjected to the GST-pulldown assay to validate its role in dmDcr-1 binding. A mutation of F419 to alanine had the strongest effect on the level of dmDcr-1 in the pulldown assay as the western blot signal for dmDcr-1 almost decreased completely (Figure 2.16 A, lane 4), while GST-Loqs-dsRBD3 efficiently precipitates dmDcr-1 (Figure 2.16 A, lane 3). As a control for unspecific binding to the glutathione beads, GST alone was used, which also showed no western blot signal for dmDcr-1 (Figure 2.16 A, lane 2). F423, which is located near F419 at the beginning of β2, is only partly exposed to the surrounding solvent and has no direct influence on the binding of dmDcr-1 as shown by the F423A mutation (Figure 2.16 A, lane 5). This underlines the specificity of F419 as it might stack as a "hydrophobic finger" into a binding site located on the helicase domain of dmDcr-1.

The A444K mutant, which partly restores region 3 as in a canonical dsRBD, has also no influence on the dmDcr-1 pulldown efficiency (Figure 2.16 A, lane 6). This finding indicates that dmDcr-1 binding is not affected by the same region which is usually required for dsRNA binding in a functional dsRBD.

It has been shown that *Loqs* can express four different splice isoforms in *Drosophila* cells (Hartig et al., 2009). These variants are referred to as Loqs-PA, -PB, -PC and -PD (Introduction, Figure

1.5). All four splice variants contain the first and second dsRBDs but differ in their C-terminal region. Loqs-PD and -PC have completely different C-termini and because of that they do not possess any affinity to dmDcr-1. Loqs-PA and -PB both contain the third dsRBD but differ in the linker region between dsRBD2 and dsRBD3. Both isoforms are able to interact with dmDcr-1 but the binding affinity is dramatically increased in case of Loqs-PB, which contains the complete interdomain-linker region between dsRBD2 and dsRBD3 (Förstemann et al., 2005; Ye et al., 2007). To further identify the amino acids, which are responsible for this difference, truncations of GST-tagged Loqs constructs were made starting from the interdomain linker full length construct (Loqs-0), to the isolated dsRBD (dsRBD3), shortening the N-terminus step by step with an increment of around 10 amino acids (Figure 2.15). To directly identify amino acids that are responsible for the dmDcr-1 interaction within the sequence of the interdomain linker, the truncations were tested in the GST-pulldown assay.

The Loqs-0 construct precipitated dmDcr-1 very efficiently (Figure 2.16 B, lane 3). The same observation was made for the truncation constructs Loqs-1, -2 and -3, which have a similar affinity as the full length linker construct, as the intensity of all dmDcr-1 signals, obtained after GST-pulldown, were almost identical (Figure 2.16 B, lane 4-6). This observation encloses the interaction surface between dmDcr-1 and Loqs to the region 337-368 within the Loqs interdomain linker. A further truncation of 11 amino acids resulting in a construct starting with aa 379 however showed a strong effect on the pulldown efficiency of dmDcr-1 with a significant reduction of the western blot signal (Figure 2.16 B, lane 7).

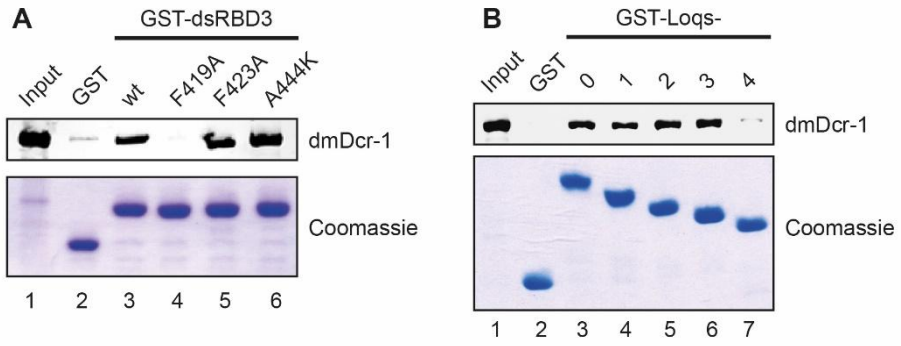


Figure 2.16: GST pulldown assay of Loqs-dsRBD3- and linker-constructs. (A) Pulldown assay of His-dmDcr1 from SF21 cell lysate with GST-Loqs-dsRBD3 and point mutants as indicated (lanes 3-6). GST alone served as control (lane 2). Bound dmDcr1 is detected by immunoblot using a specific antibody and bait proteins were stained with Coomassie blue. **(B)** Pulldown assay as in (A) using Loqs constructs Loqs-0-4 containing sequential truncations of the interdomain linker (lanes 3-7).

GST-Loqs-3, which contains amino acids 368-463 of Loqs-PB, has been identified as the minimal construct, which contains the high affinity interaction site located in the interdomain linker between dsRBD2 and dsRBD3. A secondary structure prediction of the isolated linker region, using the program Jpred4 (Drozdetskiy et al., 2015), shows major parts of this linker to be

unstructured except for the region of the minimal construct Loqs-3 where a relatively high probability for a helix was predicted (Figure 2.17). Sequence analysis of this linker region reveals an alternating sequence of hydrophobic and hydrophilic amino acids with an increment of 3-4 residues, which might form an amphipathic α-helix. The hydrophobic side of the amphipathic helix consists of L371, N373, L379/80/82 and L387.

Figure 2.17: Secondary structure prediction of the interdomain linker between dsRBD2 and 3. Potential secondary structure elements were marked as letters (H for helix) underneath the sequence of the interdomain linker. The minimal construct (Loqs-3) for the high affinity interaction with dmDcr-1 starts at residue number 368 (indicated above the sequence) within the sequence of Loqs. The last line stands for the reliability of prediction accuracy, which ranges from 0 to 9 and where high numbers indicate a strong confidence of the prediction for each respective residue.

While mutations L371A and N373E in GST-Loqs-3 had no effect on dmDcr-1 binding (Figure 2.18 A, lanes 4 and 5), there was a clear effect when L387 was mutated to alanine (See Figure 2.18 A, lane 7). The mutation of K378 to glutamate also showed reduced dmDcr-1 binding (Figure 2.18 A, lane 6). Consistently with the findings described before, the mutation F419A showed an effect also in the presence of the complete interaction surface within the minimal construct Loqs-3. The contribution from the linker region (amino acids 368-463), however, is able to partly rescue the phenotype of the F419A mutation (compare lane 4 in Figure 2.16 A with lane 8 in Figure 2.18 A). Hence, the double mutant L387A/F419A had the strongest effect in the GST-pulldown assay (Figure 2.18 A, lane 9) due to the combined mutations in the linker region and in the dsRBD3.

To further characterize the contributions of the linker- and the dsRBD3-interaction surface to the binding of dmDcr-1, GST-dsRBD3 and GST-Loqs-3 were used simultaneously in a GST-pulldown assay and the signal from the dmDcr-1 was set into relation between the two constructs (see Figure 2.18 B, lane 2 and 3). The linker containing GST-Loqs-3 construct was able to precipitate roughly 10 fold more dmDcr-1 compared to GST-dsRBD3 indicating a strong increase of the affinity to dmDcr-1 when the interdomain linker is present.

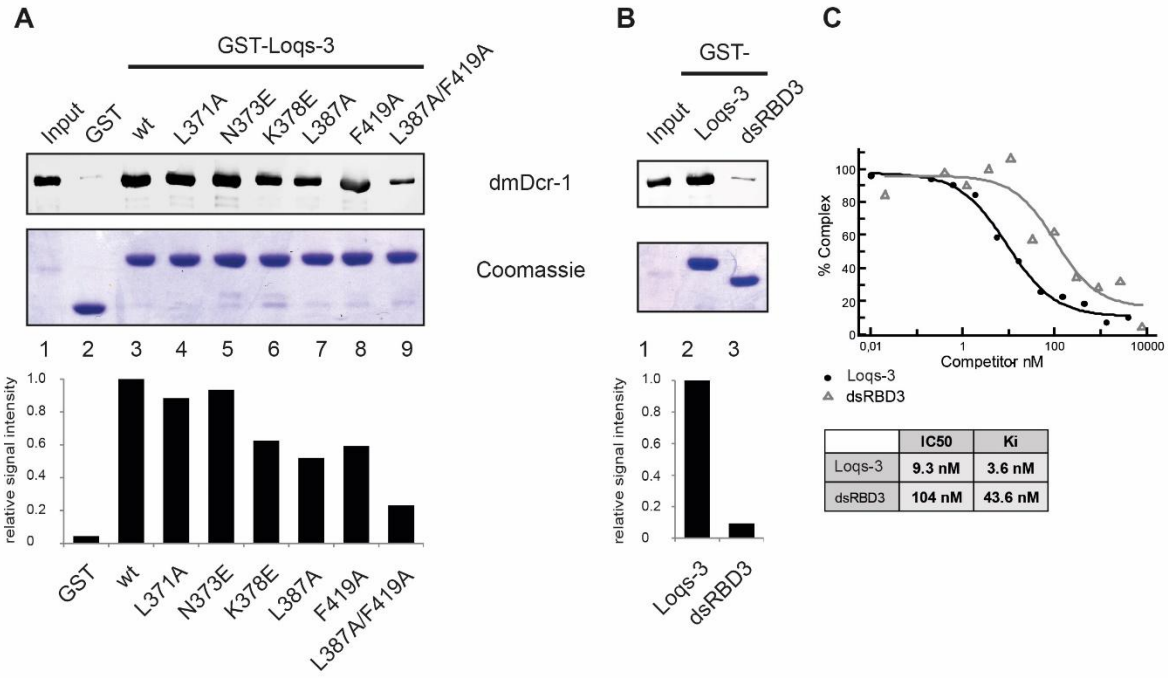
To quantify the difference of the dsRBD3 with and without the linker region in more detail, FCCS measurements were performed by our collaborators[1] (See Figure 2.1). Affinity determination was performed for GFP-labeled dmDcr1 with an Alexa Fluorophore 647 (AF647) labeled dsRBD3 (dsRBD3^{AF647}) and Loqs-3 (Loqs-3^{AF647}). For dsRBD3^{AF647} a K_d value of 191.5 nM ± 27.6 nM was

[1] See author-contribution page

measured. The Loqs-3^{AF647} construct dramatically increased the affinity resulting in a K_d value of 6.7nM ± 2.2nM (data not shown) which reflects an increase of the affinity to dmDcr-1 of around 30 fold. However both proteins were artificially modified with a fluorescent dye which could cause some deviations compared to the wt proteins.

To eliminate the influence of the fluorescent AF647 protein dye, the measurement was repeated using the unlabeled dsRBD3 and Loqs-3 as competitors against Loqs-3^{AF647} for GFP-dmDcr-1 binding. Loss of cross-correlation between the two fluorophores was measured, depending on the concentration of the added unlabeled protein (Figure 2.18, C). FCCS data provided an IC50 value of 104 nM for the dsRBD3 alone which corresponds to a K_i value of 43.6 nM. For the Loqs-3 construct an IC50 value of 9.3 nM was obtained which corresponds to a K_i value of 3.6 nM. These data exactly correlate with the relative dmDcr-1 signals from the GST-pulldown, which showed a 10 fold higher affinity when the linker is present in the bait protein.

Altogether, these observations indicate that two different binding sites in Loqs exist, which interact with dmDcr-1. The dsRBD3 has a lower affinity to dmDcr-1 but might provide specificity for the binding. The linker in contrast contributes with a high affinity to the binding but has a rather low specificity, as supported by a GST-pulldown assay whereby the linker alone fused to GST precipitated only a very low amount of dmDcr-1 but contained a strong background level of unspecific precipitated proteins (Figure 2.25 B, lane 8) suggesting misfolding and thus this control was not suitable for *in vitro* pulldown experiments.



	IC50	Ki
Loqs-3	9.3 nM	3.6 nM
dsRBD3	104 nM	43.6 nM

Figure 2.18: The interdomain linker contributes to dmDcr-1 binding. (A) GST pulldown assay for the characterization of important aas within the minimal interacting construct Loqs-3 (lanes 3-9). The signal for the precipitated dmDcr-1 was normalized to the wt construct and related to the mutant-signals. **(B)** GST pulldown assay of in presence (Loqs-3, lane 2) or absence (dsRBD3, lane 3) of the interaction surface within the interdomain linker. **(C)** Fluorescently labeled Loqs-3^{AF647} was bound to GFP-dmDcr1 and subsequently competed with different amounts of unlabeled Loqs-3 or dsRBD3. IC50 values were obtained and K_i values were calculated according to Cheng-Prusoff.

2.7 *In vivo* characterization of the Loqs-dmDcr-1 interaction interface

To examine the influence of the residues identified by the *in vitro* GST-pulldown assay, an *in vivo* assay was designed where the wildtype- and mutant-constructs of Loqs and dmDcr-1 were co-transfected into Dicer-deficient mouse embryonic fibroblasts (MEF $Dcr^{-/-}$) (Chong et al., 2010; Frohn et al., 2012; Smibert et al., 2013) in context of the respective full length proteins. As dmDcr-1 alone cannot rescue mouse Dicer loss and requires the interaction with Loqs to process mammalian miRNA precursors (Bogerd et al., 2014; Förstemann et al., 2005), the detection of rescued let-7a levels gives a functional readout for the interaction between both proteins within the cell. Therefore several mutations were introduced within the sequence of full length Loqs-PB (Figure 2.19) and the rescue of mammalian miRNA precursor processing was tested in the rescue assay by analyzing of the level from rescued mature let-7a miRNA.

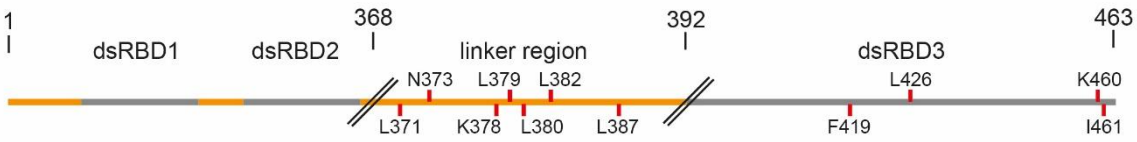


Figure 2.19: Schematic overview of full length Loqs. Mutations were marked with red lines. Double slashes indicate a change in the scale of the aa sequence.

For the rescue assay, MEF Dcr$^{-/-}$ cells were co-transfected with a modified pIRESneo eukaryotic expression vector (VP5), containing sequences of the different constructs and an N-terminal FH-tag (see Results, chapter 1.1) for anti-FLAG western blot detection. The forward transfection of the "hard to transfect" MEF Dcr$^{-/-}$ cells was quite inefficient resulting in a low number of GFP positive cells (Figure 2.20 A). However, in a reverse transfection of the MEF Dcr$^{-/-}$ cells, the efficiency could be increased to ~20 fold of additional GFP positive cells (Figure 2.20 B). Therefore all constructs used in the rescue assay were reverse transfected into MEF Dcr$^{-/-}$ cells.

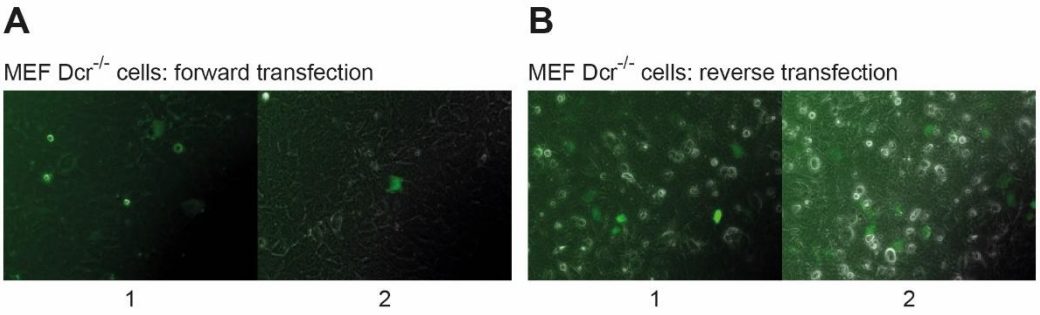


Figure 2.20: Transfection of MEF Dcr$^{-/-}$ cells. (A) Two random pictures of MEF Dcr$^{-/-}$ cells transfected with VP5 GFP by forward transfection. **(B)** Two random pictures of MEF Dcr$^{-/-}$ cells transfected with VP5 GFP by reverse transfection showing higher transfection efficiency then in (A).

After the transfection of MEF Dcr$^{-/-}$ cells with the respective proteins, total RNA was isolated for the detection of let-7a by Northern blot analysis. To confirm the right size of the rescued let-7a miRNA, total RNA from MEF Dcr$^{+/+}$ cells was loaded as a marker control (Figure 2.21 A, lane 1). As a negative control for let-7a processing and as an indicator for transfection efficiency, FH-tagged GFP was used. As expected, no signal for a mature let-7a miRNA was detectable (Figure 2.21 A, lane 2). As a second control, the human Dicer (hsDcr), which is able to efficiently process miRNA precursors also in the absence of its dsRBP partners TRBP and/or PACT (Kim et al., 2014), was expressed in MEF Dcr$^{-/-}$ cells (Figure 2.21 A, lane 3). As expected, the rescue of let-7a with the hsDcr resulted in the strongest signal for the mature miRNA. In case of the dmDcr-1 alone, on the other side, the rescue of let-7a was quite inefficient, as only a faint band was detectable for the mature let-7a miRNA (Figure 2.21 A, lane 4). Upon addition of full length Loqs-PB to the dmDcr-1 mediated rescue of let-7a processing, there was a strong increase of the

signal from the mature let-7a miRNA (Figure 2.21 A, lane 5), which was sufficient to provide a direct readout for the interaction between both proteins.

Therefore we addressed the functional relevance of the two identified Dicer-interaction modules in Loqs for miRNA processing by dmDcr1, introducing single and combined point mutations within the sequence of full length Loqs-PB. To identify differences between Loqs-PB wt and mutants, all signals from the let-7a northern blot were normalized to the wt rescue considering expression levels of each Loqs-PB construct. A mutation of L371 to alanine did not affect the interaction to dmDcr1 (Figure 2.21 A, lane 6), which is consistent with the findings from the *in vitro* binding experiments (Figure 2.18 A, lanes 3-9). Loqs N373E, however, which was able to precipitate dmDcr-1 with the same efficiency as the wt protein, rescued dmDcr-1 processing less efficiently (Figure 2.21 A, lane 7). Although the interaction is unaffected, the signal for the mature let-7a is only ~45% compared to wt, which might be due to technical differences between the two assays or between the truncated and the full length proteins. For Loqs K378E which showed a reduced binding to dmDcr-1, also the relative signal intensity of the rescue activity is decreased to ~55% (Figure 2.21 A, lane 8). Mutations of leucines 379, 380, 382 and 387, that might be part of the potential amphipathic α-helix (Figure 2.17), had the strongest effect in the rescue assay, where triple mutant Loqs L379/380/382A comprised only a weak dmDcr-1 activation with a relative signal intensity of ~20% compared to wt activation (Figure 2.21 A, lane 9) and the single point mutation Loqs L387A to even less than 20% (Figure 2.21 A, lane 10), which means almost no activation at all. These findings emphasize the importance of the hydrophobic stretch of leucines as part of a possibly amphipathic helix for the interaction to dmDcr-1. Single point mutations of this interdomain linker binding module are sufficient to efficiently disturb the interaction of Loqs-PB and dmDcr-1.

Furthermore the second binding module in the L2 loop within the dsRBD3 was investigated. Therefore a mutation of F419 to alanine was subjected to the *in vivo* and showed an effect on the processing from the let-7a pre-miRNA, where the relative signal intensity for the mature let-7a miRNA only reached ~30% of the signal from the wt rescue (Figure 2.21 A, lane 11). Consistent with the findings from the binding assay, this observation underlines the importance of the potential “hydrophobic finger” also in the context of the full length protein.

Taken these findings together, our *in vitro* binding studies and *in vivo* dmDcr-1 activity assays provided evidence that dsRBD3 harbors a low and the linker region a high-affinity dmDcr-1-binding site. Single point mutations within these modules were sufficient to disturb complex formation between Loqs and dmDcr-1 and directly affected dmDcr-1 activity in the *in vivo* rescue assay.

To further address the influence from the high affinity interaction module to dmDcr-1 binding in the full length context, an expression vector of the naturally occurring splice variant Loqs-PA was generated. Loqs-PB contains both the high affinity interaction site within the interdomain linker

and the low affinity site of the dsRBD3, the high affinity site is almost completely missing in Loqs-PA due to alternative splicing (Figure 2.21 B upper panel). Therefore, we tested both isoforms in the activity assay and consistent with the results described before, only Loqs-PB (Figure 2.21 B, lane 1) was able to process pre-let-7a efficiently whereas in Loqs-PA experiments, the level of the rescued let-7a was only slightly increased (Figure 2.21 B, lane 2) compared to a rescue with dmDcr-1 alone (Figure 2.21 A, lane 4). Therefore Loqs-PA does not only bind dmDcr-1 with reduced affinity (Förstemann et al., 2005), but was also not able to reconstitute dmDcr-1 activity in our *in vivo* assay, indicating that interaction via the linker region binding module is crucial for efficient dmDcr-1 function.

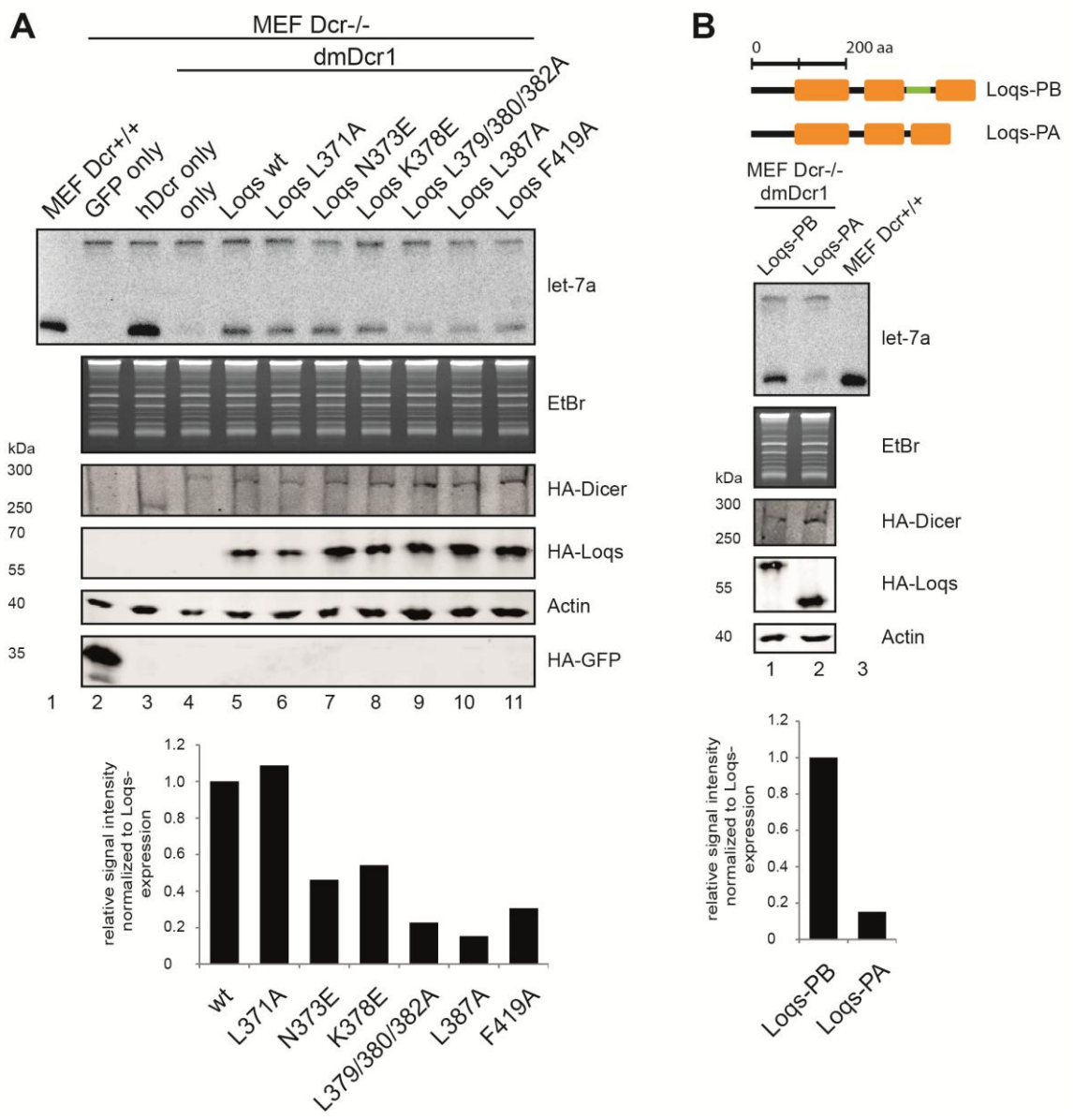


Figure 2.21: An *in vivo* Dicer rescue assays for the functional analysis of Loqs-dmDcr-1 interactions. **(A)** FH-dmDcr-1 (lanes 4-11), wt Loqs (lane 5) as well as the indicated Loqs mutants (lanes 6-11) were co-transfected into Dicer-deficient MEFs. Additionally, FH-GFP and FH-hsDcr (lanes 2 and 3) were transfected as controls. Endogenous let-7a from $Dcr^{+/+}$ MEFs (lane 1) was loaded as a marker and processing in the deficient cells was analyzed by Northern blotting (upper panel). FH-dmDcr-1 (middle panel) or FH-Loqs (lower panel) expression was analyzed by Western blotting using anti-HA antibody. Loading of the Northern blot samples was controlled by EtBr staining and loading of Western blots via an actin blot. **(B)** Comparison of the Loqs-PB (lane 2) and Loqs-PA (lane 3) isoforms in the rescue assay.

2.8 Loqs-dsRBD3 forms homodimers in the crystal and in solution

Loqs-dsRBD3 forms a tetrameric assembly in the crystal, where two monomers exchange secondary structure elements and are connected to each other (Figure 2.13 A and B). When the effect of the observed domain swaps was corrected, a dimeric assembly could be isolated. To analyze the protein contact interfaces of the isolated dimer in the crystal, a calculation was performed using the PISA-server (Krissinel and Henrick, 2007), which is an interactive tool for the exploration of macromolecular interfaces based on a structural model of a protein. The algorithm predicted a highly stable dimer in solution with a dimerization interface of about 700 $Å^2$ (Figure 2.22 A). Thereby, most of the dimerization interface is represented by hydrophobic contacts between the two monomers and a parallel orientation of aas 434-438 in each β3-strand, which forms an extended β-sheet in the middle of the dimer.

To further support our structural model of the Loqs-dsRBD3 dimer, we performed a crosslinking experiment, using disuccinimidyl suberate (DSS) with a mixture of and HA-tagged Loqs-3 and recombinant dmDcr-1. As expected, HA-Loqs-3 can be readily crosslinked to a species migrating as a dimer in a denaturing SDS-PAGE and subsequent anti-HA western blot (Figure 2.22 B). Even with increasing amounts of crosslinker, the HA-Loqs-3 dimer did not form further homomeric aggregates indicating a specific crosslink within the solution-stable dimer. Furthermore, HA-Loqs-3 was crosslinked to dmDcr-1, where a clear shift was detectable for Loqs-3-HA (green) from ~13 kDa above the dmDcr-1 band (red) at ~255 kDa (Figure 2.22 B, upper panel), indicating a native interaction between both proteins.

Also indicative for the existence of a stable dimer in solution is the running behavior of Loqs-3 in SEC (Figure 2.22 C). The column was calibrated with globular proteins of known molecular weight and respective elution volumes were indicated with arrows above the chromatograms. The wt Loqs-3 construct elutes at a relative size of ~26 kDa which is twice the MW of the monomer. As a control for the running behavior of a monomeric dsRBD, the isolated second dsRBD (dsRBD2) of Loqs was applied and elutes exactly at the predicted MW of ~14 kDa. In a further step, a mutant was designed, which was able to disturb the dimerization. Therefore L426 which is located in β2 and points towards the C-terminal helix of the second monomer was replaced against a bulky arginine which has a similar probability of occurrence in a β-sheet position (Smith et al., 1994) and thus does not disturb the formation of the secondary structure in β2. As expected, this mutant completely inhibits the formation of an intact dimer and the mutant elutes at the volume of a monomeric dsRBD (Figure 2.22 C, orange graph).

Additionally to the L426R mutant that exhibited a rather severe phenotype, amino acids K460 and I461, located in the C-terminal helix, were mutated. When K460, which is supposed to form a salt bridge to E416 (located in L2), was exchanged to glutamate, this led to an intermediate elution of the Loqs-3 mutant in the gel filtration assay (Figure 2.22 C, red graph). Unlike the L426R mutant,

K460E is still able to form dimer contacts but interactions between the two monomers were attenuated. However K460 did not form a salt bridge to E416 in the crystal which is likely an effect of the hydrophobic di-methylation of the lysine side chain which reduces the overall positive charge. Additional mutation of I461, which is located next to K460 and forms hydrophobic stacking interactions within the dimer, even reinforced the effect (Figure 2.22 C, blue graph) as the elution volume was again shifted closer to the monomeric elution volume (Figure 2.22 C, orange and green graphs).

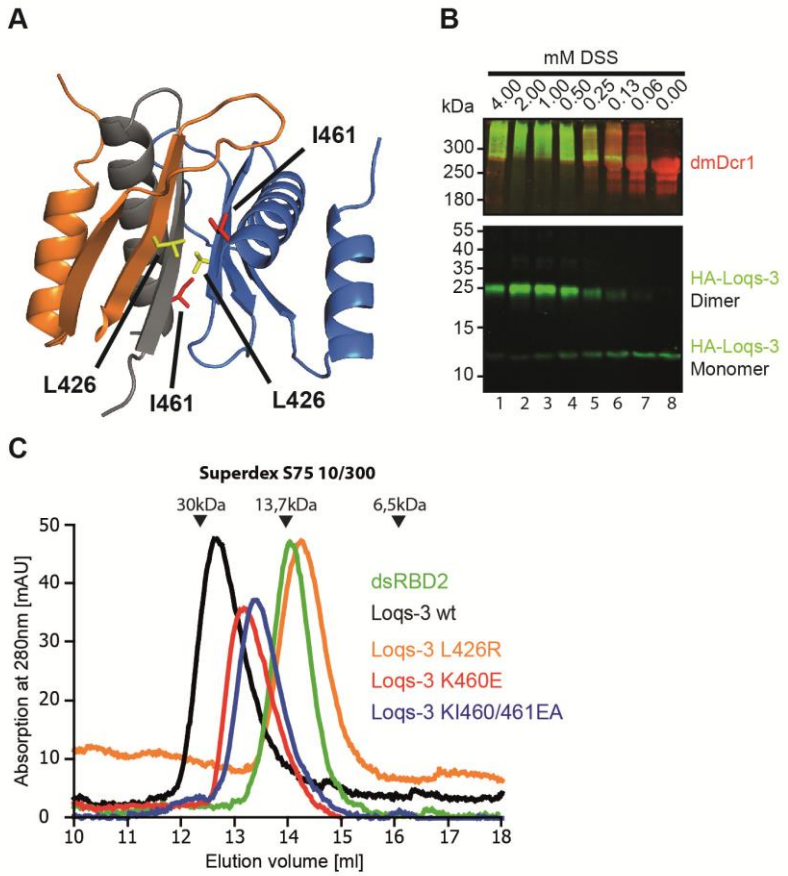


Figure 2.22: Loqs dimerizes via its dsRBD3. (A) Structure model of a dimer from native Loqs-dsRBD3 molecules. The side chain of residue L426 is shown as yellow sticks and I461 as red sticks. **(B)** Crosslinking assay of purified His-dmDcr-1 and HA-Loqs-3 (aa 368–463). Concentrations of DSS-crosslinker are indicated on top. Both proteins are detected by immunoblot, Loqs by an anti-HA-antibody (green) and dmDcr-1 by a specific antibody (red). Positions of monomeric and dimeric Loqs-3 are indicated on the right. **(C)** Size-exclusion chromatography of Loqs-dsRBD2 (green), Loqs-3 wt (black), Loqs-3 L426R (orange), Loqs-3 K460E (red), and Loqs-3 KI460/461EA (blue). Elution volumes of globular standard proteins are marked as black arrows above the chromatogram.

2.9 DmDcr-1 interaction in context of Loqs-dsRBD3 dimerization

After the identification of a stable dimer of the Loqs-dsRBD3 in the crystal and in solution we addressed the question, how the dimerization of the Loqs-dsRBD3 affects its binding to dmDcr-1. Therefore, constructs with the mutations L426R, K460E and KI460/461EA were generated for the *in vitro* binding- and for the *in vivo* activity assays, respectively. First we used the GST-pulldown assay with mutations of the minimal construct Loqs-3 (Figure 2.23 A (top panel)). Thereby, the interaction with dmDcr-1 was completely abrogated in case of L426R (Figure 2.23 A(top), lane 4). Due to the severe phenotype in the *in vitro* binding assay, L426R was subjected to an *in vivo* binding experiment with Loqs-PB full length variants in HEK 293 cells. To do so, we expressed FH-Loqs-PB wt and L426R together with myc-tagged dmDcr-1 and performed a myc-IP of dmDcr-1 (Figure 2.23 A (bottom)). The HA-western blot signal for the co-precipitated FH-Loqs-PB gave a direct readout for the interaction between both proteins in the cell. As expected, the signal for FH-Loqs-PB L426R (Figure 2.23 A(bottom), lane 3) was strongly decreased compared to the signal for FH-Loqs-PB wt (Figure 2.23 A(bottom), lane 4), indicating a strong effect of L426R on both dimerization and dmDcr-1 interaction. Consistent with that, the interaction interface of the mammalian Dicer-TRBP complex has recently been published (Wilson et al., 2015). The interface which Loqs uses for dimerization is, in case of the dsRBD3 of TRBP, bound to an α-helical insertion within the hsDcr helicase domain. Furthermore, to confirm this assumption, we subjected the L426R mutation in Loqs-PB to the dmDcr-1 activity assay. As expected, the dimerization mutant had also a reproducible effect in the dmDcr-1 activity assay, where the rescue with L426R led to no detectable increase of the signal for mature let-7a miRNA (Figure 2.23 B, lane 3) above the background level of a rescue where only dmDcr-1 was present (Figure 2.21 A, lane 4).

To further test whether Loqs-PB binds to dmDcr-1 by its dimerization surface, mutation of K460 or K460 and I461, which contribute to the dimer interface via a potential salt bridge and hydrophobic contacts and led to a destabilization of the dimer reflected by an intermediate elution in the SEC assay (Figure 2.22 C), were applied to the dmDcr-1 binding and activity assays, respectively. In TRBP, I361 participates in Dicer binding via hydrophobic interaction with L360 on Dicer. K360 on the other hand is pointing away from the interaction site suggesting no relevance for the binding to hsDcr. Consistently, in the GST-pulldown assay, GST-Loqs-3 K460E had no influence on the binding to dmDcr-1 (Figure 2.23 A (top), lane 5) and also pre-let-7a miRNA processing activity was not influenced (Figure 2.23 B, lane 4). Therefore K460 contributes rather to dimer formation then to the interaction surface with dmDcr-1. An additional mutation of I461 to alanine in the context of K460E, however, decreases the amount of precipitated dmDcr-1 in the GST-pulldown assay (Figure 2.23 A(top), lane 6) and also impairs miRNA procession activity in the rescue assay (Figure 2.23 B, lane 5).

Conclusively, residues which are responsible for the dimerization of Loqs-dsRBD3 were shown to have also an effect on the interaction with dmDcr-1. However, the results obtained so far, gave no information about the stoichiometry of both proteins in the Loqs-dmDcr-1 complex.

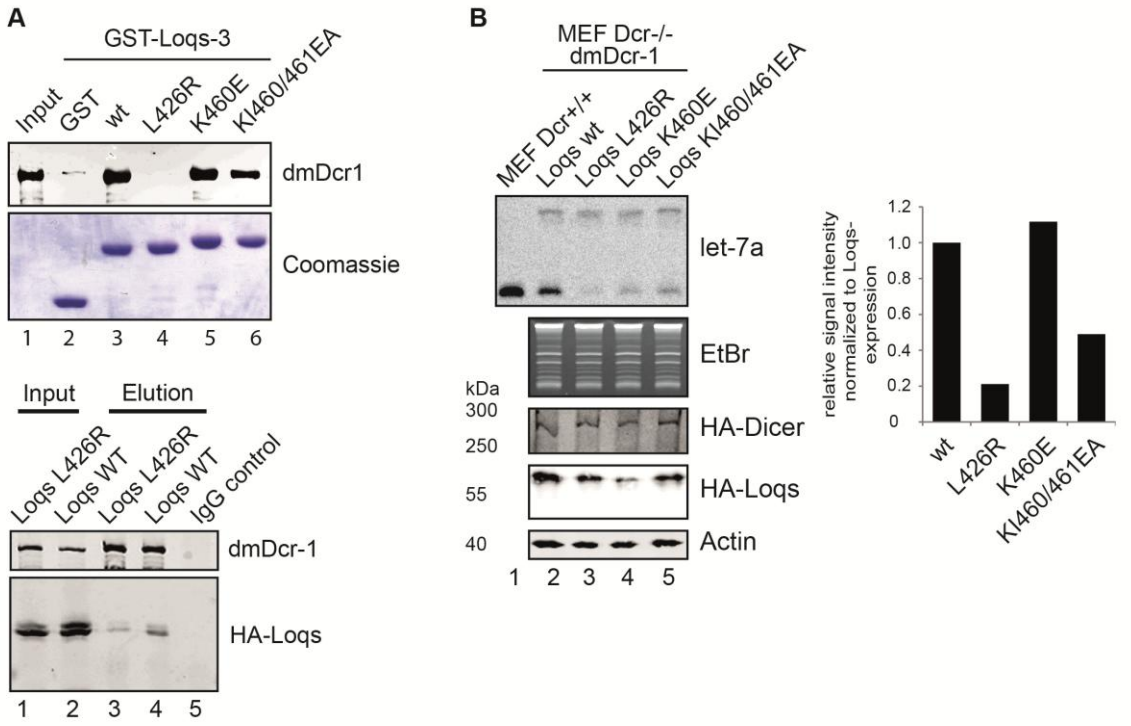


Figure 2.23: GST-pulldown and Dicer-rescue assays for the functional analysis of the Loqs-dimer. (A) (Top) GST-Pulldown assay of His-dmDcr1 with GST-Loqs-3 and mutants as indicated. (Bottom) FH-tagged wt Loqs (lanes 2 and 4) or FH-Loqs L426R (lanes 1 and 3) were co-transfected with myc-dmDcr-1 into HEK 293 cells. Myc-dmDcr1 was immunoprecipitated using anti-myc antibodies and the immunoprecipitates were analyzed by Western blotting using anti-dmDcr-1 (upper panel) or -HA antibodies (lower panel). Lane 5 shows an IgG control. **(B)** Influence of the Loqs dimerization mutants L426R, K460E, and KI460/61EA on miRNA processing.

To address the question, whether Loqs-dsRBD3 binds to dmDcr-1 as a monomer or a dimer, we used single molecule immunoprecipitation analysis[1] (Figure 2.24 A). The single molecule pull-down (SiMPull) assay that combines the principles of a common pull-down assay with single molecule fluorescence microscopy enables direct visualization of individual cellular protein complexes by immobilizing them on a quartz surface (Jain et al. 2011). This method is highly sensitive and can determine the stoichiometry of biomolecular complexes. Therefore we used the recombinant GFP-tagged dmDcr1 (See chapter 2.6) and Loqs-3^{AF647}, that contained exactly one single AF647 label within the dsRBD3. A preformed complex of GFP-dmDcr-1 and Loqs-3^{AF647} was immobilized on a passivated quartz slide, using a biotinylated GFP antibody (Figure 2.24 A) and fluorophores were exited using total internal reflection fluorescence (TIRF) microscopy.

[1] See author-contribution page

Single AF647 fluorescence spots were detectable in the dmDcr1 containing fraction and were mostly absent, when GFP-dmDcr-1 was omitted (Figure 2.24 B, C), indicating a relatively weak unspecific binding of Loqs-3^{AF647} to the surface and providing evidence to the specificity of the immobilization strategy. The photosystem of single fluorescent emitters is not permanently stable upon excitation and shows an abrupt decrease in fluorescence intensity, a process also known as photobleaching, which then can further be used to count the number of molecules present at an individual fluorescent spot. For the Loqs-3^{AF647}/GFP-dmDcr-1 complex, one fraction (~90%) showed one single bleaching step and a minor fraction (~10%) showed two bleaching steps (Figure 2.24 D, E). This could rather be explained by an unspecific background binding of the labeled dimer then by a dimer bound to GFP-dmDcr-1. In summary these findings indicate that, although Loqs forms dimers in solution, it is bound as a monomer to dmDcr-1, which is consistent with the findings for the hsDcr-TRBP complex (Wilson et al., 2015).

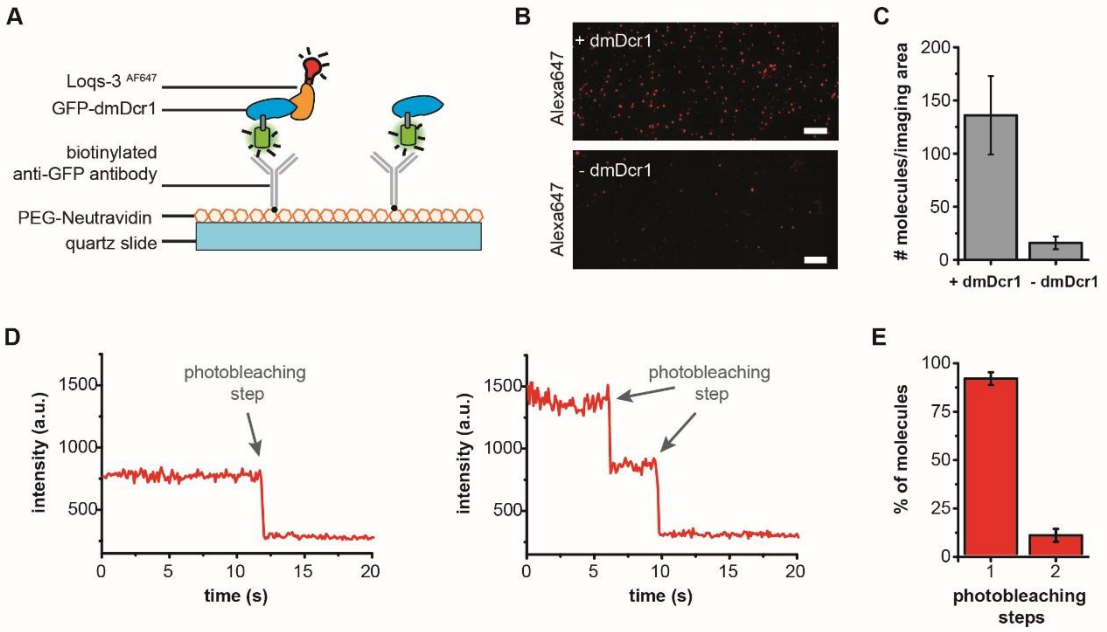


Figure 2.24: Single-molecule co-IP assay of the Loqs-dmDcr-1 complex. (A) Schematic overview of the assay **(B)** Total internal reflection fluorescence (TIRF) images of Alexa647-labeled molecules (Loqs-3^{AF647}) pulled down via GFP-dmDcr-1. The GFP-dmDcr-1 fusion protein was immobilized on the quartz slide via a biotinylated anti-GFP antibody (scale bar: 10 µm). As a control, Loqs was flushed over the slide without prior incubation with GFP-dmDcr-1. **(C)** Average numbers of Loqs-3^{AF647} molecules per imaging area (6800 µm2). **(D)** Representative single-molecule fluorescence transients of Loqs-3^{AF647} molecules that exhibit one-step and two-step photo bleaching. **(E)** Photo bleaching step distribution for Loqs-3^{AF647} bound to dmDcr1 indicating a 1:1 stoichiometry for the Dicer–Loqs complex. Data were collected from four independent experiments from a total number of 1558 molecules.

2.10 Characterization of the TRBP-hsDcr interface for the analysis of post translational modifications on hsDcr

For the characterization of the Loqs-dmDcr-1 interface, an *in vitro* binding assay was established (see chapter 2.6). Furthermore, this assay was adapted for the Loqs homolog TRBP, to characterize the binding between TRBP and hsDcr. To do so, we fused GST to different truncations of the interdomain linker between the dsRBD2 and dsRBD3 (aa positions 231-291) together with the dsRBD3. Interestingly one part of the linker region (aa 261-291) is required for the solubility of the dsRBD3 construct and truncations within this region led to insoluble proteins (data not shown), suggesting a requirement of residues within this region for the stability of TRBP in solution.

In an initial GST-pulldown screen, the binding of the TRBP constructs, TRBP-1 (aa 231-363), TRBP-2 (aa 241-363), TRBP-3 (aa 261-363) and the interdomain linker alone (aa 231-291) to overexpressed FH-hsDcr were investigated (Figure 2.25 A). As expected, all constructs containing the interdomain linker connecting dsRBD2 and 3 were able to efficiently pulldown FH-hsDcr (Figure 2.25 A, lanes 10-12) except the linker alone or the GST-control (Figure 2.25 A, lanes 2 and 13). To test whether Loqs is able to interact with the hsDcr, truncations of Loqs were used in the binding assay to precipitate FH-hsDcr. However, all Loqs constructs used for the pulldown, only had a very weak affinity to FH-hsDcr (Figure 2.25 A, lanes 3-8). The interdomain linker of Loquacious (Figure 2.25 A, lane 8) comprised the highest level of precipitated FH-hsDcr which might be due unspecific binding caused by misfolding in the absence of the dsRBDs (see chapter 2.6).

Interestingly, when dmDcr-1 was applied to the same a GST-pulldown assay, all truncations of TRBP (Figure 2.25 B, lanes 10-12), except the isolated interdomain linker between dsRBD2 and 3 (Figure 2.25 B, lane 13), were able to efficiently precipitate dmDcr-1, in an equal or even higher level than the Loqs constructs (Figure 2.25 B, lanes 3-7). Taken these findings together there is some evidence that Loqs harbors a special feature, like its tendency to form dimers in solution (see chapter 2.8), to specifically recognize dmDcr-1 and not the human Dicer. As hsDcr has a higher sequence similarity to dmDcr-2 (data not shown) this feature might therefore be important for the discrimination between the two *Drosophila* Dicers.

Due to the high pulldown efficiency of the TRBP dsRBD3, which was not only restricted to hsDcr, the binding assay was used for the enrichment of the hsDcr for MS analysis. Therefore TRBP-3 was used to precipitate either the endogenous or overexpressed FH-tagged human Dicer from HEK293 cells (Figure 2.25 C). The precipitation efficiency was assayed by western blotting using a Dicer-specific antibody. Additionally a western blot against Ago2 was performed to investigate if human Ago2 is co-precipitated in the GST-pulldown. As Ago2 is part also of the RLC, it is likely to be co-precipitated within the pulldown experiment and would provide

information about the phosphorylation status that is present in proteins within the RLC. From a lysate, containing the overexpressed FH-hsDcr, the pulldown worked quite efficiently with a strong signal for hsDcr (Figure 2.25 C, lane 6). The endogenous hsDcr was precipitated less efficiently but there was still a prominent Dicer band (Figure 2.25 C, lane 7), that looked promising for MS analysis. However, there was no band detectable for Ago2, that co-precipitated with the hsDcr in the GST-TRBP-3 pulldown and also in the Ago2 control-IP (Figure 2.25 C, lane 3) there was no detectable band for hsDcr, indicating that a possible complex of Ago2 and hsDcr might not be stable under tested conditions. Thus MS analysis was performed on the precipitated endogenous hsDcr and overexpressed FH-hsDcr. As MS analysis required high amounts of proteins, a large scale GST-pulldown, using GST-TRBP-3, of either endogenous or overexpressed hsDcr was performed and loaded onto an 8 % SDS-gel which was stained with Coomassie (Figure 2.25, D).

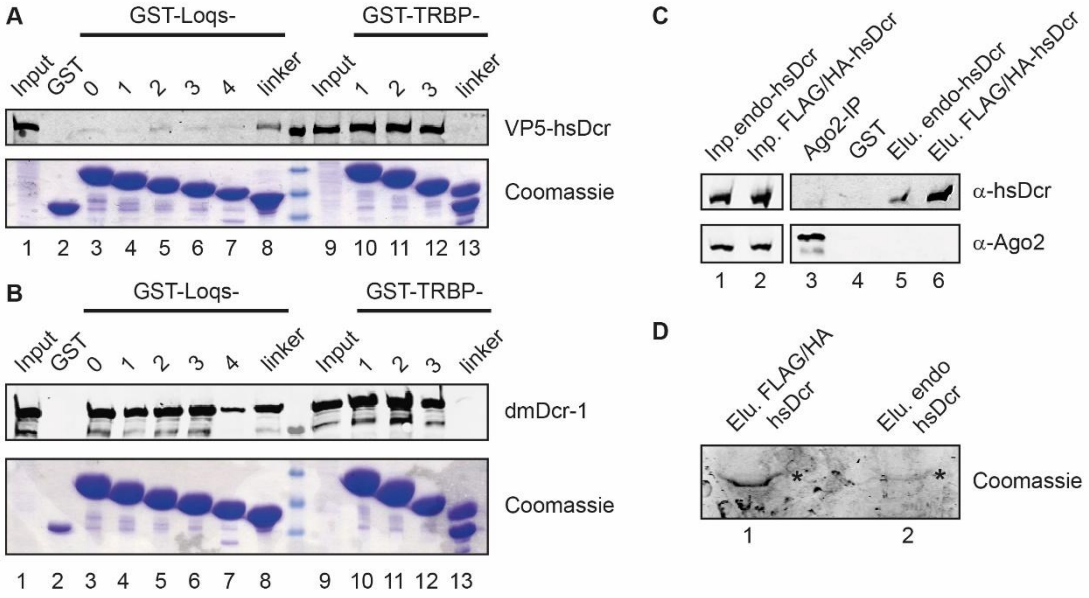


Figure 2.25: GST-pulldown of the hsDcr. (A) GST-pulldown assay of overexpressed hsDcr with Loqs- and TRBP-linker constructs. **(B)** GST-pulldown assay of overexpressed dmDcr-1 with Loqs- and TRBP-linker constructs. **(C)** Pulldown of endogenous and overexpressed FLAG/HA-tagged hsDcr with TRBP-3 as bait protein (upper panel). Control for co-precipitation of endogenous Ago2 with an Ago2 specific western blot (lower panel). **(D)** Coomassie gel of a large scale GST-pulldown for endogenous and FLAG/HA-tagged hsDcr1. Bands used for MS analysis were marked with an asterisk.

The bands containing hsDcr were excised from the gel and used for MS analysis after elution and tryptic digestion. Digestion with trypsin, which cleaves after lysines, generates a specific cleavage pattern with unique peptide composition for any protein. These peptides can be separated according to molecular weight and their overall charge and detected on the MS detector which results in their amino acid composition. Post-translational modifications alter the molecular weight and charge of peptides and therefore can be differentiated from unmodified peptides. For

this study, the digested hsDcr samples were analyzed for post translational modifications with the focus on phosphorylation sites. Thereby both samples comprised high sequence coverages with 62.5 % for the endogenous and 56.3 % for the FH-hsDcr.

In case of the endogenous hsDcr, only one phospho-site, Y1330 (Figure 2.26, bold), which is located in the RNase III domain, could be identified. This is probably due to the lower amount of precipitated protein (Figure 2.25 D). For the overexpressed FH-hsDcr, more protein was available for the MS-analysis and six additional phospho-sites were identified. Four were located within functional domains, Y153 in the helicase domain, S678 in the dsRBD, S1016 in the PAZ domain and S1470 in the RNase III domain of hsDcr (Figure 2.26). The remaining two sites were the adjacent S1252 and S1255 and are located in an interdomain region between the PAZ and RNase III domain (Figure 2.26). Phospho-sites S1016, S1252, S1330 and S1470 were already annotated on the PhosphoSitePlus® (www.phosphosite.org) databank. Y153, S678 and S1255, however, have not been annotated yet from any post-translational modification screen.

Table 2.3 gives an overview of the type of amino acid, the respective position and the score of the identified phospho-sites within the hsDcr sequence. Sites of endogenous hsDcr are marked with bold letters.

Taken these findings together, a fragment of the human TRBP (TRBP-3) can be used for an efficient hsDcr-isolation. The amount of precipitated hsDcr was sufficient to identify several phospho-sites within the sequence of the hsDcr. Furthermore, TRBP-3 was able to efficiently precipitate the dmDcr-1 and therefore might also be used for the isolation and characterization of Dicer-proteins from other species.

To sum up the first part of this thesis, novel insights into the miRNA biogenesis pathway and RISC-loading process in flies and humans were obtained. Using X-ray crystallography the structure of Loqs-dsRBD3 was solved, providing evidence of an interaction surface that is either bound by dmDcr-1 as a monomer or dimerizes in the absence of dmDcr-1. Additionally, several residues within the interdomain linker region between dsRBD2 and 3 in Loqs-PB were identified which provide a high affinity dmDcr-1 binding site. Whereas Loqs-dsRBD3 comprised features that allowed a specific recognition of dmDcr-1, the dsRBD3 of TRBP was able to efficiently pulldown both, the dmDcr-1 and hsDcr and therefore might be used for the enrichment of Dicer proteins from different species in general. From the enriched hsDcr several phospho-sites were obtained by MS.

After the processing of pre-miRNAs by the Dicer-dsRBP complex, the mature miRNAs were loaded onto an Ago protein. Therefore an Ago-loading assay using FCCS was established, resulting in a detailed kinetic observation of the loading process from different human wt and mutant-Ago proteins.

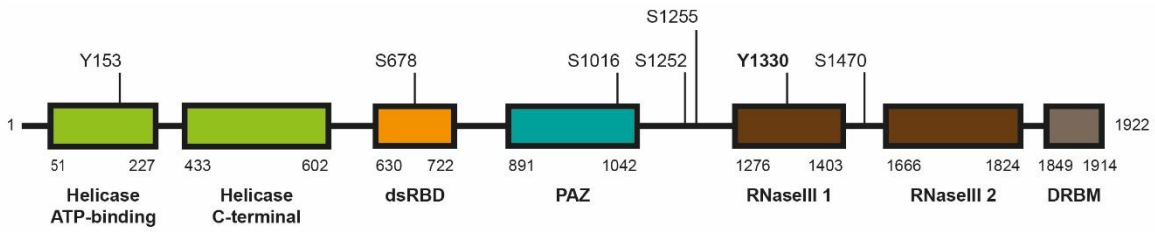


Figure 2.26: Schematic overview of the identified phospho-sites of the human Dicer

Table 2.3: Identified phospho-sites of the human Dicer

hsDicer	aa	position	score	location
endogenous	Tyrosine (Y)	**1330**	19.6	RNase III 1
overexpressed (FLAG/HA)	Tyrosine (Y)	153	31.9	Helicase (ATP-binding)
	Serine (S)	678	32.7	dsRBD
	Serine (S)	1016	14.2	PAZ domain
	Serine (S)	1252	24.3	Interdomain region (PAZ/RNase III 1)
	Serine (S)	1255	15.9	Interdomain region (PAZ/RNase III 1)
	Serine (S)	1470	71.9	Interdomain region (RNase III 1+2)

In the miRNA biogenesis pathway, several RNA-binding proteins, that contain unique compositions of RNA-binding domains, are required for efficient substrate recognition and processing, to generate a mature miRNA, which allows an Ago protein to function as a post-transcriptional regulator. Loqs thereby contains three dsRBDs, where two dsRBDs (dsRBD1+2) are able to specifically interact with the dsRNA-stem regions of pre-miRNA. Ago on the other hand is a RNA binding protein that binds to ssRNA by its PAZ and Piwi domains. The NHL-domain of *Drosophila* Brat belongs to another family of RNA binding proteins, which is involved in the post transcriptional regulation of many mRNA targets that contain a specific Brat binding motive. In the second part of this thesis, the molecular details of the target-recognition of the Brat-NHL domain were structurally characterized through crystallization of the Brat-NHL domain in complex with its consensus motif RNA.

3. Crystallization of the Brat-NHL domain in complex with its consensus motif RNA

In previous studies, the NHL domain of Brat has been identified to contact the *hunchback* mRNA in a sequence-specific manner, which leads to translational repression (Loedige et al., 2012) and (Loedige et al., 2014). In follow-up experiments, that were performed by others in our lab, the consensus motif with the sequence 5'-UUGUUG-3' was identified and showed a significantly high affinity for the Brat-NHL domain. To identify the molecular details behind the interaction between the Brat-NHL domain and its consensus front motif RNA, the complex was formed, isolated and applied to X-ray crystallography.

3.1 Purification of the Brat-NHL-RNA complex

For the generation of the Brat-NHL-RNA complex, the highly purified NHL domain of Brat (aa 756–1037) as described in (Loedige et al., 2014) was used together with a 15 nucleotide long RNA, containing the consensus motif within the first 6 nucleotides of the sequence (consensus front RNA) followed by a stretch of 9 Us (5'-UUGUUGUUUUUUUUU-3'). The consensus motif was identified with multiple expectation and maximization for motif elicitation (MEME) (Bailey and Elkan, 1994) and a second method termed RNAcompete (Ray et al., 2009), in which the recombinant NHL domain of Brat together with a complex pool of short RNAs was used *in vitro*, to screen for high affinity sequence motifs.

To analyze the complex formation, the Brat-NHL domain and the consensus front RNA, which were both provided by postdocs in the lab[1], were loaded separately onto a Sepharose 75 10/300 gel filtration column. The highly purified Brat-NHL domain eluted at 13.2 ml as a homogenous peak (Figure 2.27 A). The consensus front RNA however eluted at approximately 14.6 ml but contained a contaminant nucleic acid species, which eluted in front of the consensus front RNA peak at ~13.5 ml (Figure 2.27 B). This contamination might derive from a degradation of the hammerhead ribozyme system (Price et al., 1995), used for generating the RNA, as it increased after longer incubation for the autocatalytic cleavage (data not shown). If an equal amount of Brat-NHL is added to the consensus front RNA, a clear shift of the complex could be detected to ~12.2 ml (Figure 2.27 C). The contaminant remained unaffected referred to its elution behavior (peak at 13.5 ml) upon addition of the Brat-NHL domain.

[1] See author-contribution page

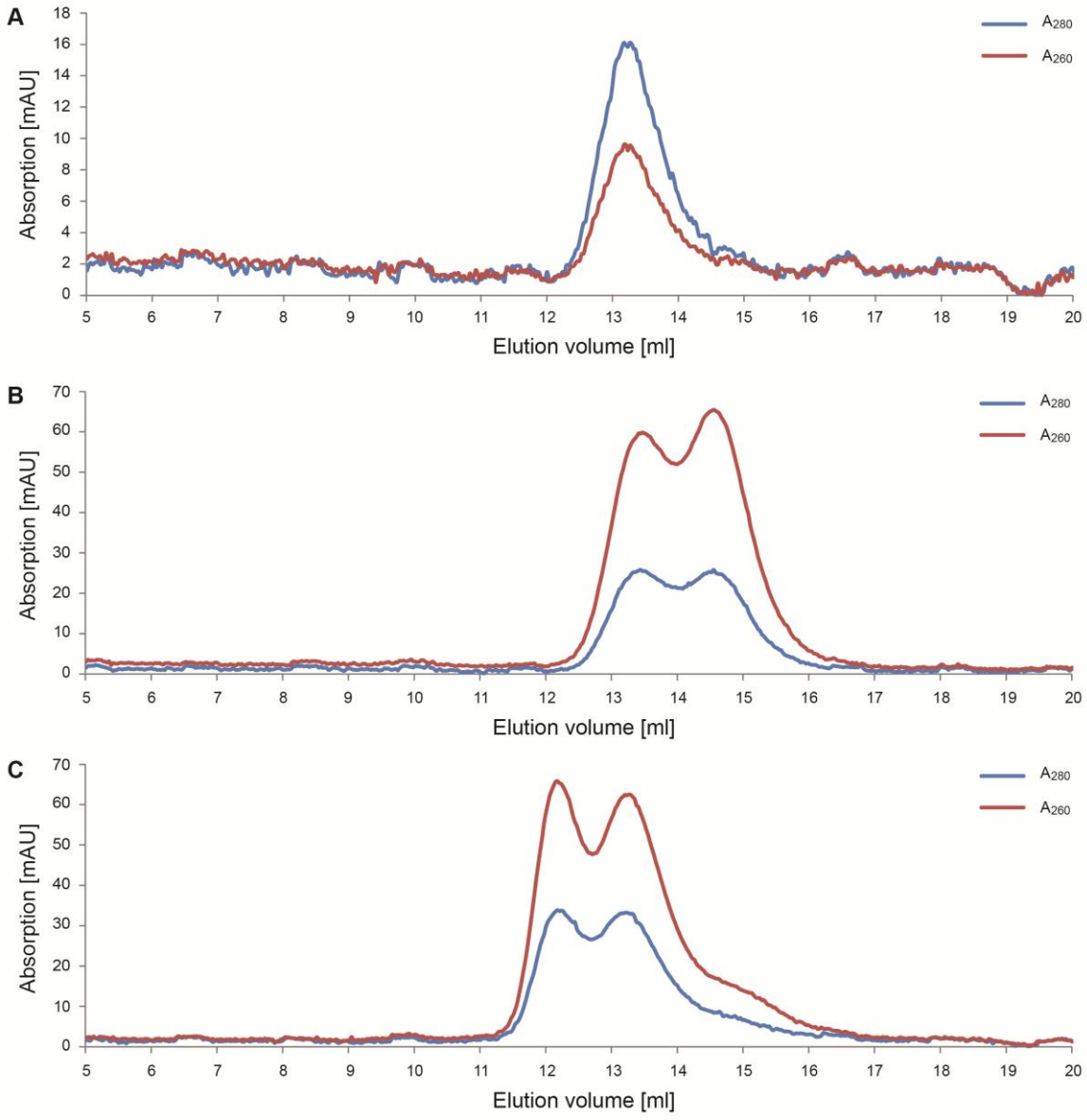


Figure 2.27: Analytical gel filtration of the Brat-NHL-RNA complex. (A) Analytical gel filtration of the Brat-NHL domain alone. Absorption of the protein at 260 nm (red line) and 280 nm (blue line) was monitored. **(B)** Analytical gel filtration of the consensus front RNA alone. **(C)** Analytical gel filtration of the Brat-NHL domain in complex with the consensus front RNA.

The complex preparation for crystal screening was scaled up and Brat-NHL was supplemented with an equal amount of consensus front RNA. The amount of RNA was adjusted according to the ratio between contamination and consensus front RNA. Fractions containing the complex were isolated after a first Sepharose 75 10/300 gel filtration column (Figure 2.28 A) and concentrated. A second gel filtration showed only a slight shoulder deriving from the contamination and a single peak for the purified and highly stable complex (Figure 2.28 B). Complex containing fractions were pooled, concentrated to ~2.3 mg/mL and used for crystallization studies.

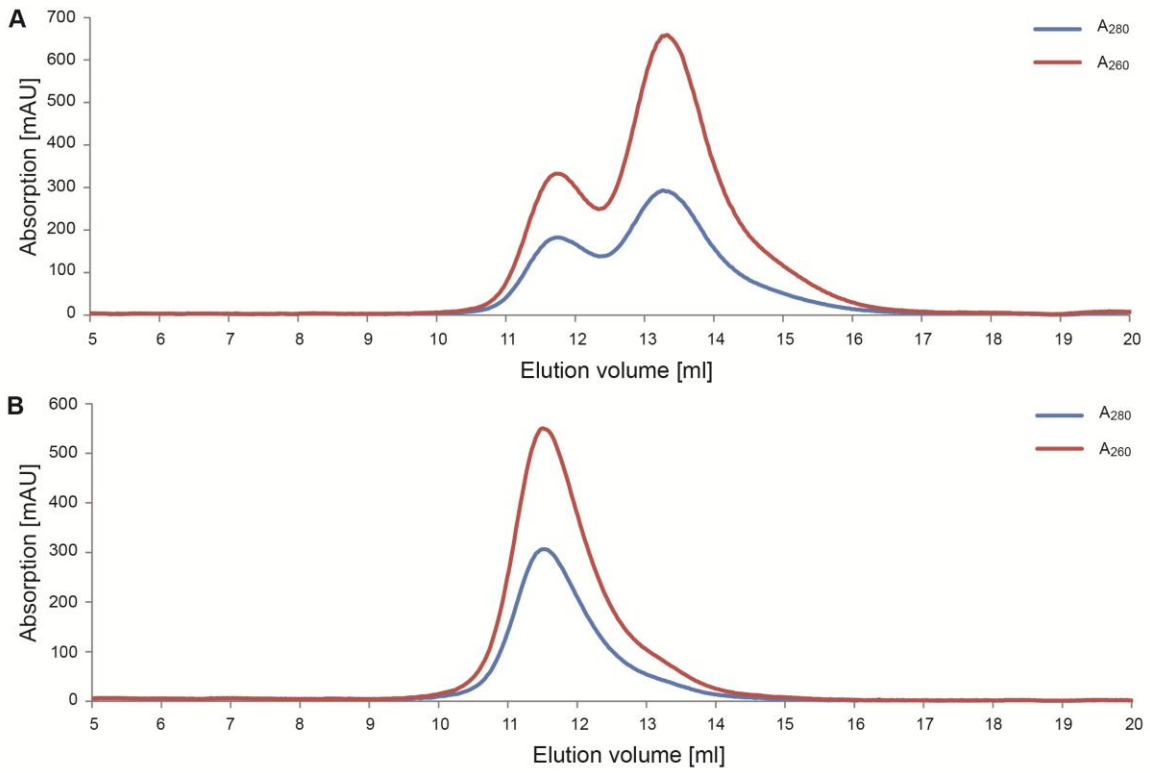


Figure 2.28: Preparative gel filtration of the Brat-NHL-RNA complex. (A) First preparative gel filtration of the Brat-NHL-RNA complex (~12 ml). An RNA contamination included in the consensus front RNA sample could also be detected (~13.5 ml). **(B)** The complex containing fractions from the first gel filtration were concentrated and subjected to a second gel filtration column. The contamination was strongly reduced and the complex remained stable.

3.2 Crystallization of the complex

In an initial crystallization trial, the complex was subjected to the commercially available crystal screens JBScreen Classic 1-8, JBScreen Wizard I-IV and Natrix I and II (Table 4.11) in 96 well sitting drop plates (see methods, chapter 7.1). The complex crystallized after one day in a single condition containing 100 mM MES pH 6.5, 2 M ammonium sulfate, 5 % (w/v) PEG 400 at 20°C. The crystals grew as a tetragonal bipyramids with a size of less than 0.1 mm.

For X-ray crystallographic measurements, conditions with initial crystals, in most cases, require further refinement to improve the quality of the crystals. Therefore refinement screens, using the "hanging drop" and the "sitting drop" method in 24 well crystallization plates, were performed, where the drop size, the pH, the concentration of the precipitant and from other components of the initial condition were altered. However, within these refinement screens, the crystals from the initial condition were hardly reproducible and only formed in two additional conditions (Data not shown). The crystals within these conditions did not differ from the initial crystals and thus, as refinement screening failed to optimize the crystals for X-ray crystallographic measurements, crystals from the initial sparse matrix screen were used. Therefore, crystals were either briefly soaked in cryo-buffer (20% glycerol, 5% PEG400 and 500 mM ammonium sulfate) or directly

taken out of the condition without further treatment and frozen in liquid nitrogen and stored for X-ray crystallographic measurements.

3.3 Data collection and structure determination

The data collection was performed at the BESSY (Berlin) on the beamline BL14.2 where cryo-buffer treated and untreated crystals were measured. The crystals, which were treated with cryo-buffer, diffracted with a resolution range of ~3 Å in the highest resolution shell. The untreated crystals, however, showed a much better diffraction down to ~2.7-2.3 Å in the highest resolution shell. The reservoir solution from the initial sparse matrix screen was already sufficient to act as a cryo-buffer and was able to protect the crystals upon freezing as no ice rings were observed on the diffraction images of the dataset. Ice rings develop upon too early nucleation of ice crystals in the crystal-surrounding solution and therefore act as an indicator for insufficient cryo-protection of the crystal.

Datasets were recorded from untreated crystals. The Brat-NHL-RNA complex crystallized in the tetragonal space group $P4_122$ with a resolution of 2.3 Å. Table 2.4 shows some statistical parameters of the collected data.

Table 2.4: Data collection statistics

	native
Wavelength (Å)	0.9814
Resolution range (Å)	19.77-2.3 (2.4-2.3)
Space group	P 4_1 2 2
Unit cell, a b c, α β χ	70.19 70.19 279.87, 90 90 90
Total reflections	388255(46244)
Unique reflections	32238 (3769)
Multiplicity	12 (12.3)
Completeness (%)	99.99 (100.00)
Mean I/sigma(I)	11.77 (2.42)
Wilson B-factor [$Å^2$]	26.99
CC½ [%]	99.5 (74.2)

Values as reported by *CORRECT*. Values in parentheses refer to the highest resolution shells.

Phasing of the Brat-NHL-RNA crystal was performed using molecular replacement with the program Phaser (McCoy, 2007). The crystal structure from the apo Brat-NHL domain (pdb entry 1Q7F) was used as a search model (Edwards et al., 2003). Two monomers of Brat-NHL with high Z-scores were placed into the asymmetric unit cell (Figure 2.29 A). The Z-score judges the quality

of the *in silico* solution of the phase problem and values above 8, which was the case for Brat-NHL, imply an unambiguous solution.

Model building was performed in COOT (Emsley et al., 2010) and refinement of the model against the density was performed with PHENIX.*refine* (Afonine et al., 2012) using simulated annealing and non-crystallographic symmetry restraints (NCS). If the refinement was performed without consideration of hydrogen atoms, there was a gap between R_{work} and R_{free} of more than five percent, which indicates a strong model bias, which derives from the search model that was used for molecular replacement. The refinement under consideration of the hydrogen atoms resulted in a much lower gap with a R_{work} value of 0.188 and R_{free} value of 0.219 (Table 2.5) and is therefore more reliable.

Two monomers of the Brat-NHL domain and one molecule of the consensus front RNA were built into in the asymmetric unit cell, where the Brat-NHL monomers were bound to one RNA molecule simultaneously (Figure 2.29 A).

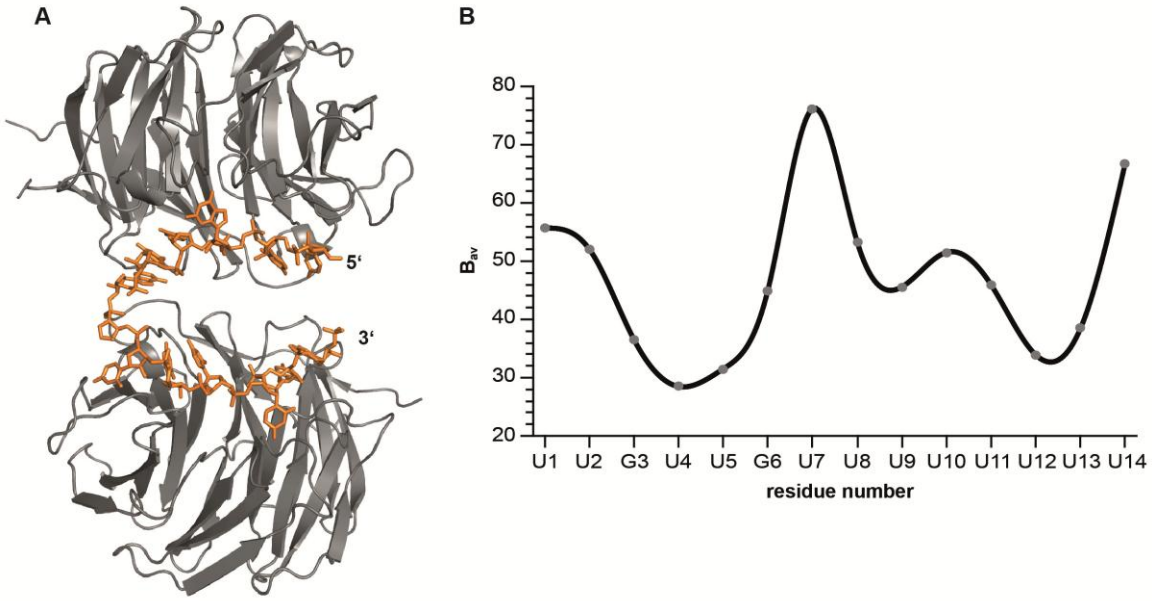


Figure 2.29: Asymmetric unit cell of the Brat-NHL-RNA complex. (A) Overview of the asymmetric unit cell where two Brat-NHL domains bind to one molecule of the consensus front RNA. **(B)** B-factor plot of single nucleotides from the consensus front RNA. High B_{av} values indicate a strong flexibility and weak binding whereas low values indicate less flexibility and strong binding to the protein backbone.

For an estimation of temperature-caused vibrational motion of different parts of the structure, the B-factors were used. Thereby, atoms with low B-factors can be assigned to parts of the structure that is well ordered, whereas atoms with large B-factors generally belong to parts of the structure that are flexible. The B-factor of the protein chains was 29.67 $Å^2$ for chain A and 31.85 $Å^2$ for chain B indicating a relatively rigid and well-ordered binding platform for the RNA. Only in a

small loop region of chain A, there was only poor electron density and therefore S825 could not be built into in the structural model.

Regarding the bound consensus front RNA on the other side, the distribution of the B-factors varies strongly between each single nucleotide position (Figure 2.29 B). A plot of the B-factors for each nucleotide against its position within the RNA sequence clearly highlighted the strongest binding or the lowest flexibility, respectively. The minima of the B-factor plot occurred at positions G3, U4 and U5 in monomer A and positions U11, U12 and U13 in monomer B. Nucleotide U7, which is located between the monomers showed the highest flexibility and the pyrimidine base therefore was not resolved in the electron density. U15 at the 3' end of the consensus front RNA could also not be built because of poor electron density, which was additionally split to two possible conformations. All building attempts to place the terminal nucleotide into the electron density, lead to an increase of the R_{work} and R_{free} values and therefore U15 was omitted from the structural model. Model statistics and structure refinement data are given in Table 2.5.

Table 2.5: Model statistics and structure refinement

Parameter	**Brat-NHL/RNA complex**
R_{work}/R_{free}	0.188/0.219
Modelled polypeptides Chain A/B	758-824,826-1036 / 758-1035
Modeled RNA Chain C	Nts 1-14, 15 not resolved
No. of non-hydrogen atoms macromolecules water/ligands	5133 4697 338/98
RMS(bonds) [Å] RMS(angles) [°]	0.003 0.75
Ramachandran favored (%) Ramachandran outliers (%) Clashscore	96.7 0 1.28
Average B-factor [$Å^2$] Protein (Chain A/B) RNA ligands/solvent	31.72 29.67/31.85 46.27 43.90/33.50

3.4 The crystal structure of the Brat NHL-RNA complex

To illustrate the molecular details behind this sequence specificity, the crystal structure of the Brat-NHL domain in complex with a 15 nt long consensus front RNA has been solved at a resolution of 2.3 Å (Figure 2.29 A). The Brat-NHL domain assembles as a six-bladed β-propeller, which surrounds a solvent filled channel at the center of the domain. The top surface of the NHL-domain contains mainly positively charged amino acids, which are mostly forming a platform for the interaction with the backbone of nucleic acids. This conformation remains mainly unchanged upon substrate binding. The bottom surface, however, mainly has a negative overall charge and was not involved in RNA binding in the crystal structure. The binding of the consensus front RNA is mediated via three different binding regions, which are responsible for specific base recognition (Figure 2.30 B). U1 and U2 stack against each other and are bound between blades II and III of the β-propeller. G3 is bound between blades IV and V flipped out from the π-stacking interaction with the residual RNA strand. The last three nucleotides from the consensus motif U4, U5 and G6 stack against each other and were coordinated between blade VI and I (Figure 2.30 A). The pyrimidine base of U1 from the consensus motif interacts with the protein backbone via π-stacking to Y829 and the base forms two hydrogen bonds from O2 to R847 and one hydrogen bond from N3 to N800. However, the electron density was weak for the base of U1 and also the high B-factors indicate some flexibility within this region. N2 from the second U is coordinated by E782 and the base stacks against U1 (Figure 2.30 C, panel 1).

The 5’ phosphate of G3 is bound by K891 and R875, which also contact the 5’ phosphate of the neighboring U4 and thus enabling a conformation whereby G3 can flip out of the π-stacked RNA strand to form specific contacts with the protein backbone. Thereby, the aromatic ring of G3 is stacked onto F916 which enables the purine base to form two hydrogen bonds with peptide bonds of the protein main chain. Ring position N1 forms a hydrogen bond with the peptide carbonyl group of E915 and O6 with the peptide amine of K891 (Figure 2.30 C, panel 2). The rigid assembly of the protein backbone specifically enables a guanosine to form a maximum of possible interactions explaining the binding preference for a G at position 3 within the consensus motif.

Nucleotide U4 and the adjacent U5 form most of the interactions with the Brat-NHL domain and were tightly bound to the protein backbone. This is reflected by the minima of the B-factor plot (Figure 2.29 B). The 2’OH of the ribose attached to U4 forms hydrogen bonds to the side chains of N933 and R934 and therefore represent contacts that are specific for an RNA template, as DNA, on the other side, lacks this 2’OH group. Ring positions O2 and O4 of the uracil base from U4 make specific hydrogen bonds to the side chain amide of N933 and the peptide N-H of C1004. In addition, N3 forms a hydrogen bond to a water molecule that is coordinated by the protein backbone via the peptide carbonyl of Y959 and the peptide N-H of N976. U5 stacks to the base of U4 and contacts the protein backbone via hydrogen bonds of ring positions N3 to the peptide

carbonyl of N976 and O4 to its side chain amide. The coordinated water, which binds to U4, also contacts U5 at ring position O2 (Figure 2.30 C, panel 3).

G6 is the last nucleotide of the consensus motif and extends stacking interactions with U4 and U5. N1 binds to the peptide carbonyl of H1001 and N2, which is specific for guanosine is also able to make hydrogen bonds to the backbone carbonyls of H1001 and N978 (Figure 2.30 C, panel 4).

The RNA, which was used for the crystallization contained the consensus motif within the first six positions followed by a stretch of nine uracil bases. This stretch is bound by a second Brat-NHL domain allowing a direct comparison of the complete versus an incomplete binding motif. U9 and U10, which were equivalent with U1 and U2, show the same coordination in both monomers. U11, which is equivalent with G3, is also flipped out of the RNA strand and stacks against F916. However the pyrimidine base is only able to form one single hydrogen bond to the backbone N-H of K891, which might explain why poly(U) also exhibits a relatively strong affinity to the Brat-NHL domain (data not shown). U12 and U13, which are equivalent to U4 and U5, also show the same coordination in both monomers. The G6 equivalent U14 is not able to for similar hydrogen bonds as the guanine base and therefore the electron density is only weakly defined at this position.

Taken these findings together, the crystal structure of the Brat-NHL domain bound to its RNA consensus binding motive reveals how sequence specificity is brought about. Thereby, three pairs of neighboring β-blades from the β-propeller form three specific binding clefts, which allow a perfect accommodation of the six bases from the consensus motif through a network of base-specific hydrogen bonds.

Within this thesis, novel insights into Dicer processing and Ago loading concerning the miRNA biogenesis pathway could be obtained. Furthermore the molecular details behind the interaction of a consensus RNA motif and the Brat-NHL domain could be identified at atomic resolution. In the following section these results will be discussed under consideration of the recent scientific background.

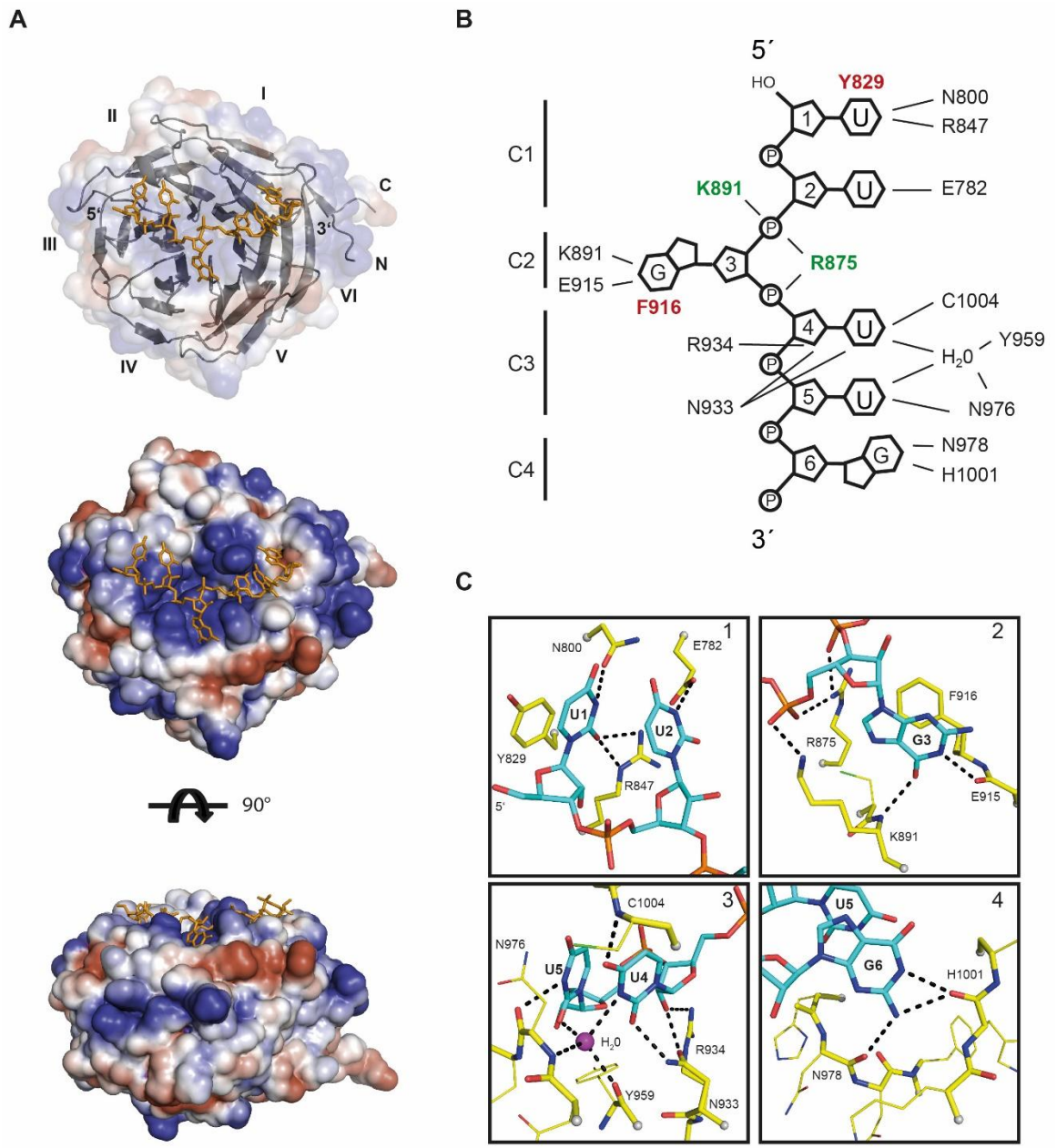


Figure 2.30: Structure of the Brat-NHL domain bound to its consensus motif RNA. (A) Structure model of the Brat-NHL domain bound to RNA. Top: model of the six-bladed β-propeller of the NHL domain viewed from the top surface. The individual blades are numbered, and the bound RNA is shown as a stick representation. Center: electrostatic surface representation of the Brat-NHL top side with bound RNA. Bottom: side view of the complex with the electrostatic potential shown for the protein domain. **(B)** Schematic representation of the protein/RNA contacts. Side chains forming stacking interactions to RNA bases are shown in red. Ionic interactions with the phosphate backbone of the RNA are indicated in green. Hydrogen bonds with the bases or sugars of the RNA are shown in black. **(C)** Close-up view of the RNA binding pockets. RNA is shown in cyan, and protein residues are shown in yellow. Hydrogen bonds are marked by dashed lines. Side chains not interacting with the RNA are drawn as thin lines, and interacting side chains and backbone atoms are shown as thick lines and labeled. All chain cuts are marked with halos.

III. Discussion

In eukaryotic cells, gene expression is highly regulated by processes, which require a large set of different RBPs, of which each displays specific RNA- and protein-protein interactions. The interaction with target RNA can be either mediated directly by RBP-RNA interactions or indirectly via base paring with an RBP-bound complementary guide RNA strand. In this thesis, X-ray crystallography and single molecule spectroscopy have been used for the characterization of representatives from both groups.

MiRNAs are small RNAs of 21-24 nts in length and are endogenously expressed as pri-miRNAs by PolII in the nuclei of eukaryotic cells. Two RNase III enzymes, nuclear Drosha and cytosolic Dicer, are responsible for the generation of mature miRNAs including their correct ends. During the last several years, many novel insights into the mechanisms underlying the processing of miRNAs could be obtained. However, lots of molecular details of the interaction between Dicer and its dsRBP partners during pre-miRNA processing as well as strand selection during Ago loading are still unclear. Therefore, one aim of this thesis was to structurally and functionally characterize interactions between Dicer and associated dsRBPs as well as further investigating determinants of strand-specific Ago-loading processes following pre-miRNA processing.

TRIM proteins belong to one of the largest protein families and are conserved among many different species. Thereby, TRIM family members containing the NHL-domain have been identified as key regulators of developmental transitions. In progenitor stem cells, for example, TRIM family members promote differentiation, while inhibiting cell growth and proliferation. The NHL-domain of the Drosophila TRIM-NHL protein Brat has recently been identified as a novel RNA-binding domain (Loedige et al., 2014) classifying TRIM-NHL proteins as direct RBPs. Therefore, the aim of the second part of this thesis was to crystallize the NHL domain of Brat together with its consensus RNA motif, to elucidate how binding specify and selectivity is achieved.

1. RISC loading in mammals

The mammalian RISC loading follows a cytosolic procession step where Dicer, in complex with its dsRBP partners, removes the stem loop from the pre-miRNA and produces the second end of the mature miRNA. Thereby, the miRNA duplex has a 2 nt overhang at both 3’ ends and a phosphate group at each 5’ end. Regardless of the sequence of each strand, the overall shape of the duplex is symmetric and the ends are chemically equal. One strand of the duplex, however, is recognized as the guide strand and is incorporated into an Ago protein while the passenger strand is rapidly degraded. About the loading of the duplex strand and the determination of the correct incorporation of the guide strand little is known so far. In fact, the small RNA duplex itself

already represents one determinant for the selection of the guide. Base pairing at both ends of the duplex strongly contributes to guide strand selection. Following the so called asymmetry rule, the strand with the thermodynamically less stably paired 5' end is preferentially loaded into Ago proteins and becomes the guide strand (Khvorova et al., 2003; Schwarz et al., 2003).

For knockdown experiments of any gene of interest in the cell it is therefore important to design siRNA duplexes with an effective guide strand loading into Ago. The asymmetry rule thereby needs to be considered. Until today neither sequence specificity nor any influence of the siRNA structure on guide incorporation into Ago could be shown. X-ray crystal structures of human Ago2 in complex with a single-stranded guide RNA revealed sequence-independent interactions between amino acids and guide RNA backbone phosphates (Elkayam et al., 2012; Schirle and MacRae, 2012). Further X-ray crystal structures of ternary hAgo2-guide-target complexes revealed a binding pocket for position 1 of the target RNA which specifically binds adenosine (Schirle et al., 2015, 2014).

Based on these observations, it seems that additionally to the thermodynamic asymmetry, at least one part of the sequence of the guide strand could be important for the efficiency of the guide strand incorporation. Therefore, one aim of this thesis was to analyze sequence-specific effects of siRNA duplexes on loading efficiencies. Thus an FCCS assay was designed to monitor binding of Ago to a set of different guide RNAs.

1.1 Ago loading resolved by FCCS

FCCS was chosen to monitor Ago-loading processes, because it is a highly sensitive method detecting molecular interactions on a nanomolar level (Schwille et al., 1997). One other advantage of this assay is its application in a high-throughput scale allowing screening of various siRNAs in parallel. To enable detection and characterization of the binding partners, both species need to be fluorescently labeled. Therefore, Ago was fused to an N-terminal GFP-tag, whereas the siRNAs were synthesized with a Cy5 label at the 3' end of the predicted guide strand. To circumvent variability of different transient transfections, stable HEK TREx-293 cell lines were generated using a vector encoding GFP-tagged Ago protein.

To detect Ago-loading events, a lysate of respective GFP-tagged Ago proteins was generated using hypertonic swelling of the cells (Dignam et al., 1983b). Unfortunately, freezing of the lysate resulted in the formation of aggregates, which interfered with the FCCS measurements. Addition of supplements according to Martinez et al. (2002) stabilized the lysate enabling freezing without a major loss of activity. The amount of glycerol was adjusted from 10 % to 5 % because higher concentrations of glycerol (>10%) alter the refraction index of the solution and thus hampers the detection of the fluorescently labeled particles within the measurement chamber. To remove Ago containing aggregates and higher MW complexes (Figure 2.2, A), centrifugation of the lysate at

100.000 x g was necessary because these particles caused inhomogeneity in the fluctuation curve of the FCCS measurements.

Introducing all these modifications, a lysate was generated that enabled the observation of binding events between a Cy5 labeled siRNA and active GFP-Ago protein, more or less uninfluenced by the fluorescent labels (Figure 2.2, B,C and D). Furthermore the lysate was stable for several freeze-thaw cycles without the loss of Ago loading activity.

1.2 siRNA-loading into Ago proteins

Following the optimization of a binding competent lysate for the FCCS assay, measurements with 3' Cy5 labelled siRNAs were performed. These measurements are essential to allow further large scale screening attempts with different unlabeled siRNAs and to validate whether the efficiency of a siRNA-mediated knockdown is influenced by its affinity to an Ago protein.

Based on a functional knockdown screen that was performed by Michael Hannus (siTOOLs Biotech GmbH), where several siRNAs against PLK were tested, two siRNA candidates were randomly chosen. For the FCCS assays, both carried a Cy5 label at their 3' ends. For both siRNAs the correlation between knockdown efficiency and incorporation efficiency into Ago was determined. The K_d value of the less effective siRNA (si856, K_d ~544 ± 1 nM) is about fivefold increased, compared to the more efficient siRNA (si428, K_d ~126 ± 15 nM). To exclude variations in the lysate quality, both siRNAs were measured in a lysate from the same batch. Both siRNAs display thermodynamic asymmetry but whereas in case of si428 the first base pair is an U/A pair which is less stable base pairing than the 5' G/C pair of si856. This difference in the 5'-end explains the less efficient binding of si856 to Ago2. Compared to literature the K_d values of the binding from Ago2 to a guide 3' Cy5 labeled 21-mer siRNA duplex were slightly increased compared to 21-mer siRNA duplex with a guide strand that is labeled with fluorescein at position 14 (counted from the 5' end) with a K_d value of ~48 nM (Deerberg et al., 2013). This difference of the K_d values might derive from an influence of the fluorescent label on the binding of the 3' end of the guide strand at the PAZ domain of Ago2 but still lies in a rage of a certain error between both methods of measurement.

Due to this potential effect of the Cy5 label on the 3' end binding of the siRNA guide strand and to circumvent high costs for the synthesis of labeled siRNA strands, a competition assay with the already characterized Cy5 labelled si428 was established, which can further be applied to any unlabeled siRNA. Thereby the K_d values of the unlabeled inhibitor siRNAs, the K_i value, can be obtained. Notably it was essential to make serial dilutions of the unlabeled siRNA duplex, containing constant concentration of the labeled si428 duplex, prior adding the mixture to the lysate. If the lysate was incubated with the labeled si428 before the addition of the unlabeled siRNA, no competition of si428 could be observed due to a tight binding of the labeled strand.

This strong binding affinity might be caused by a processing of the siRNA duplex by Ago2 within the cellular lysate. After the passenger strand is removed, the single stranded guide strand comprises a higher affinity to Ago then the duplex strand (Deerberg et al., 2013). This effect has also been shown by other groups, which tried to compete endogenously loaded siRNAs in *Drosophila* Ago2 upon addition of an exogenous let-7 siRNA (Iwasaki et al., 2015). Thus the competition of loaded Ago with a siRNA duplex is strongly impaired and a pre-incubation with the labelled siRNA duplex is not advisable for the competition assay.

When the competitor siRNA duplex was added simultaneously with the labelled si428, a competition of both duplexes for the binding to Ago2 could be measured resulting in K_i values, which represent the K_d of the competitor. A set of 8 unlabeled siRNAs against different members of the import family was used. All siRNAs showed an efficient knockdown activity except for one pair of siRNAs designed against the *Transportin-1* (Tnpo) mRNA. As the K_d values for siRNAs with high knockdown efficiencies distributed of ~30 nM, the Tnpo si1, which had a quite inefficient knockdown efficiency, comprised a remarkably and reproducible shift of its K_d value to ~76 nM. This observation might be indicative that the affinity of the Tnpo si2 to Ago2 is decreased and that this correlates with the decrease of the knockdown efficiency.

For characterization of the data obtained from the FCCS screen, further experiments should be performed to collect more details of the Ago loading process. First, all siRNAs which were used in this assay were added to the lysate without a 5' phosphate. This phosphate has been shown to be critical for efficient Ago loading (Elkayam et al., 2012; Ma et al., 2005; Nykänen et al., 2001; Schirle and MacRae, 2012). It has been shown, that siRNAs are phosphorylated by cellular kinases upon addition to cellular lysates, however this process and further downstream loading processes require ATP. To exclude deviant measured K_d and K_i values, these experiments should be repeated with siRNA duplexes that contain a phosphate group at their 5'ends.

Second, the FCCS assay as conducted can only resolve the binding of siRNA duplexes to Ago, whereas further downstream events like passenger strand removal and/or cleavage (Leuschner et al., 2006; Matranga et al., 2005; Rand et al., 2005) are hidden from view. Therefore a GFP-Ago-IP should be performed after loading of the siRNA duplex under conditions of the FCCS assay. Northern blot analysis of the RNA extraction would further provide evidence which one of the two strands is incorporated as the guide and which one is removed as the passenger after loading or if both strands are detectable, which would be indicative for an incomplete unwinding of the duplex. Alternatively, a double labeled siRNA with a second fluorescent label at the passenger strand could also provide information about the fate of the siRNA duplex, once it is bound to an Ago protein.

1.3 Loading of human Ago 1-4

Besides GFP-Ago2, other stable cell lines expressing GFP-Ago1, GFP-Ago2 H807R a slicer inactive mutant of Ago2, GFP-Ago3 and GFP-Ago4 were generated. Lysates from each cell line were titrated with the Cy5 labeled si426 and K_d values were obtained using FCCS. Disregarding of the kinetics from each loading event to the respective Ago proteins, loading could be observed in any tested lysate. However, with different K_d values for each Ago protein. For Ago1, the weakest binding affinity was measured. For Ago2 and the slicer deficient mutant Ago2 H807R, almost identical K_d values were measured. For Ago3 the strongest and for Ago4 the second strongest affinity could be measured.

The monoclonal cell line of GFP-Ago3 and the polyclonal line of GFP-Ago4 expressed each respective protein less efficiently after induction with tetracycline and yielded in a lower amount of overexpressed protein after lysate preparation. Therefore, the effect of a higher affinity to the si428 duplex might rather derive from a better availability of other Ago endogenous loading factors like Hsp90, FKBP 4/5, Dicer and TRBP, which were included in the total cellular lysate, then from protein specific properties, which promote a stronger binding of si428. Additionally, lysates from high expressing cells were diluted with lysate from uninduced HEK-293 cells to adjust equal particle counts for the measurements of all Ago proteins. The lysate from the uninduced HEK-293 however was produced earlier and might therefore stronger be depleted from ATP which has been shown to be important for siRNA 5' phosphorylation and unwinding of the duplex for RISC activation (Nykänen et al., 2001).

To exclude an effect from the quality of different lysates, new lysates should be generated and measured directly after lysis. Additionally the expression from all GFP-Ago proteins should be adapted, that all constructs were equally expressed at the time point of cell lysis. Additionally, ATP should be added to the lysis buffer and measurements should be repeated with phosphorylated 5' ends of the siRNAs as already mentioned above.

1.4 Ago 2 phosphorylation

Ago proteins are specialized RNA binding enzymes and play a key role in small RNA-mediated gene silencing. About their regulation, however, little is known up to date. MS data from endogenous proteomic screens or from overexpressed Ago proteins identified several potential phospho-sites (Rüdel et al., 2011; Sharma et al., 2014) but for most identified phospho-sites, no functional relevance has been assigned.

The FCCS assay is a powerful tool to obtain information about the binding behavior of Ago to a dsRNA substrate. Thereby, the binding of any GFP-tagged mutant that mimics a potential phosphorylation within the sequence of Ago to a CY5-tagged siRNA can be monitored. To control the functionality of the assay in general, mutants of the putative phospho-site Y529, which

have been characterized earlier (Rüdel et al., 2011) have been generated with an N-terminal GFP tag and mutations of Tyr to Ala, Phe, Gln and Glu. Consistent with the results obtained in the paper, a mutation of Tyr to Glu had the strongest effect on the loading of Ago2 (Figure 2.6, B). As in the case of a potential phosphorylated Y529, the side chain from the Glu mutant sterically blocks the entry of the siRNA duplex into the 5' binding pocket, which is additionally enforced by repulsion of a negative charge from the Glu carboxyl group and the also negatively charged 5' phosphate. Furthermore, the Ago 2 Y529E mutant can be used as a negative control for the FCCS assay because the binding of small RNA duplexes is strongly impaired (Figure 2.6, B) as no correlation between GFP-Ago2 and the Cy5-tagged dsRNA could be measured even in conditions with high concentrations of the labeled siRNA.

A second mutant, where Y529 is exchanged to Gln, shows a similar phenotype in the FCCS assay as the Glu mutant, however a weak residual binding affinity could be measured. This observation is consistent with the findings in Rüdel et al., (2011), where this mutant, together with the mutants Y529A and F, were still loaded with endogenous miRNAs after FH-Ago2 FLAG-IP and subsequent northern blot analysis and which was not shown in the case of Y529E. The Gln side chain has identical space-filling properties as the Glu side chain and is able to sterically block the entry of the 5' phosphate into the 5' binding pocket, however the effect of a strong electrostatic repulsion is reduced in the polar, uncharged side chain of Gln. This might explain the observed shift of the K_d value of the Y529Q mutant (Figure 2.6, D).

A mutation of Y529 to Ala, compared to Gln and Glu, does not block the 5' binding pocket and allows an entry of the 5' phosphate into the 5' binding pocket of Ago2. In the wildtype protein, the first nucleotide of the siRNA guide strand forms π-stacking interactions with Y529, which stabilizes a conformation of the first nucleotide in a kinked out conformation. The base stacking with the second nucleotide of the guide strand thereby is interrupted. In the Y529A mutation, the 5' binding pocket is accessible for the siRNA strand, however there are no π-stacking of the first base which could stabilize an optimal 5' nucleotide binding. Therefore this might explain the shift of the K_d value to ~1.5 µM which is indicative for a low total affinity of the respective Ago2 mutant to the dsRNA duplex (Figure 2.6, D).

Consistent with the observations described above, a mutation of Y529 to Phe has the closest similarity to the binding behavior of the wt protein. The aromatic ring of the Phe side chain is able to stack against the first base of the siRNA and allows an optimal binding of the first nucleotide of the siRNA guide strand. Therefore the K_d value of the Y529F mutant is almost identical with the respective K_d of the wildtype Ago2 (Figure 2.6, D). Taken these findings together, the stacking interaction of the first nucleotide to Ago2 seems to be crucial for an efficient binding of the siRNA guide strand.

From a MS screen performed with overexpressed FH-Ago2, S798 could be identified to be phosphorylated in the respective overexpressed background. Within this study, the mutants

S798A and S798E were applied for cleavage assay and controlled for loading with an endogenous miR-19b (Rüdel, 2010). For a Glu substitution, the cleavage activity was strikingly reduced compared to an Ala substitution. This effect did not derive from an altered guide strand binding behavior as both mutants were still equally loaded with endogenous miR-19b. Due to a lack of structural information at the time of this study, this effect could not be explained in detail. However after the determination of the structure from human Ago2 (Elkayam et al., 2012; Schirle and MacRae, 2012), it has been shown, that S798 is directly involved in the binding of miRNA by a hydrogen bond with the phosphate backbone from nucleotide 5. To examine whether mutations S798A and E of GFP Ago2 have an impact on the binding of Cy5 si428, lysates from both mutants were measured in the FCCS assay. The FCCS assay showed no detectable difference of the K_d values from both mutants (S798A and S798E) and both K_d values were identical to the one which was measured for the wildtype protein (Figure 2.6, D).

The putative phospho-site S798 is located inside a flexible loop region ranging from residues 787-800 within the PIWI domain ~10 Å away from to the catalytic DDH motif. An electrostatic repulsion caused by a mutation of Ser to Glu, which mimics the phosphorylated state from the sidechain of S798 might induce a movement of the flexible loop towards the catalytic core of Ago2. A movement in other directions is blocked by the protein backbone (Figure 3.1). This movement might further cause a disturbance within the catalytic DDH motif or simply interfere with the orientation of the target strand prior cleavage could explain effects which are observed in the thesis of Rüdel, (2010). The affinity to a siRNA strand however is generally unaffected by the S798E mutation, indicating an overall intact binding pocket for the RNA duplex and a relatively weak contribution of the hydrogen bond from S798 to the backbone of a siRNA. Therefore a phosphorylation of S798 might be required for regulation of Ago2 in general or in other Ago related processes which are independent of Ago slicer activity, e.g. during dsDNA break repair (Gao et al., 2014; Wei et al., 2012) or during Ago/RNA-mediated interaction with chromosomal DNA (Hall et al., 2002; Janowski et al., 2006), where the slicer activity might be disturbing and only RNA binding activity is needed. However further experiments would be required to clarify the mechanistic details behind this speculation.

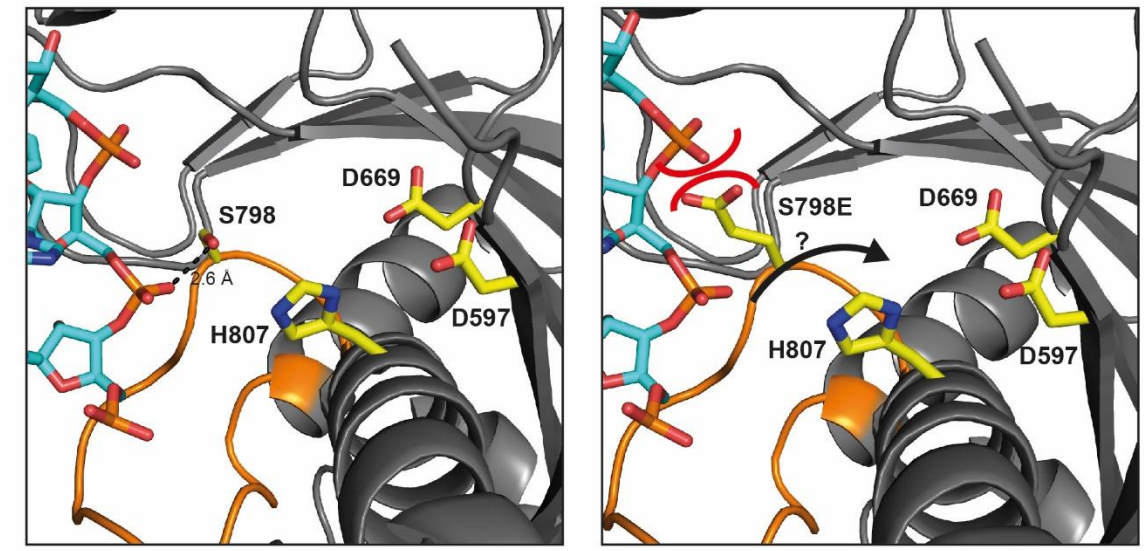


Figure 3.1: Phosphorylation of S798 in human Ago 2. Left: The PIWI domain of Ago2 (PDB-ID: 4F3T) is shown in grey. S798 is located in a flexible linker region (orange) which is in close proximity (~10 Å) of the catalytic DDH motive. Amino acid side chains were shown as yellow sticks. Right: Mutation of S798 to E or a potential phosphorylation of the endogenous Ser causes sterical and electrostatic repulsion (red braces) at this position within the flexible loop. This repulsion might cause a shift of the loop towards the catalytic center (black arrow), which then further interferes with the catalytic slicer activity of Ago2.

Taken these findings together, the FCCS assay is a powerful tool to study binding of a labeled dsRNA duplex to GFP-Ago in solution on a single molecule level. It is likely that co-localized particles refer to GFP-Ago and the Cy5 labeled single strand, however, control experiments (e.g. northern blot or binding study with a dual labeled siRNA duplex) still need to be performed. All GFP tagged Ago proteins were integrated into stable inducible cell lines. Therefore other potential candidates, which were implicated with Ago loading, could be either overexpressed or knocked down and the FCCS assay. This would provide a direct readout for potential effects on Ago loading. Furthermore a protocol needs to be established for the FCCS assay, where cells can be transiently transfected instead of generating stable cell lines, as this process is rather time consuming and not applicable for screening of high amounts of different mutants.

2. Processing of miRNAs by a dsRBP/Dicer complex

In flies, the processing of pre-miRNAs into mature miRNAs requires a functional complex of an RNase III enzyme which collaborates in complex with a dsRBP. In *Drosophila*, two different RNase III enzymes, namely dmDcr-1 and dmDcr-2, are responsible for the generation of small ncRNAs from long double-stranded precursor species. Thereby, dmDcr-1 in complex with Loqs-PB and to some minor extend with Loqs-PA is able to process pre-miRNAs (Förstemann et al., 2005). In humans, there exists only one Dicer (hDcr), which interacts with the Loquacious homologs TRBP and PACT. Disruption of the complex in both species has been shown to

interfere with the processing of pre-miRNA substrates as well as for a subset of miRNAs strand selection processes (Fukunaga et al., 2012; Wilson et al., 2015). However for mammalian Dicer, processing of pre-miRNAs in general seems to be rather independent of its dsRBP-partners, however the processing of several iso-miRs, which derive from the same pre-miRNA but were alternatively cleaved by Dicer, resulting in distinct seed regions, is affected by a complete loss of the human dsRBPs PACT and TRBP (Kim et al., 2014). Furthermore, in the absence of TRBP and/or PACT, strand selection by Ago2, where Ago2 is loaded with the guide strand and the passenger strand is removed and degraded, seems not to be altered in any of the miRNAs identified by small RNA sequencing (Kim et al., 2014).

In *Drosophila*, the processing of miRNAs, however, is dependent of a functional complex of dmDcr-1 and Loqs (Bogerd et al., 2014; Fukunaga et al., 2012). So far, little is known about the molecular details of the dmDcr-1-Loqs interaction. For the characterization of the molecular details of this interaction, the crystal structure of Loqs-dsRBD3 has been solved at 2.65 Å resolution in this thesis. Furthermore, two different assays have been established to identify critical residues within the interaction surface by a screen of different point mutations *in vitro* and *in vivo*.

2.1 Purification, crystallization and structure determination of Loqs-dsRBD3

For the crystallization of Loqs-dsRBD3, folded secondary structure elements of the dsRBD3 have been identified by *in silico* secondary structure prediction (PredictProtein). After the rigid core of the dsRBD3 has been identified and flexible unfolded regions at both sides of the dsRBD3 were omitted, a construct that contained residues 392-463 of Loqs-PB was recombinantly overexpressed. Purification of the protein yielded in a high grade of purity, which was required for X-ray crystallography. Subsequent crystallization screens with the wt protein failed to identify conditions, which resulted in initial crystals. Therefore the solubility of the construct was altered by methylation of Lysines, which lead to an increase of the hydrophobicity of the protein. Only one condition was identified, which resulted in crystals with an impaired crystal growth behavior. Regardless of the crystal growth behavior, the structure of Loqs-dsRBD3 could be solved comprising an arrangement of two domain swapped monomers packed against up to 4 single monomers within the asymmetric unit cell. This uncommon arrangement is supposed to originate from lacking crystal contacts and therefore might explain the macroscopic observation of the crystal growth behavior. The domain swap of two monomers might rather be required to compensate for lacking crystal contacts then having a functional meaning of this assembly *in vivo*.

2.2 The dmDcr-1-Loqs-dsRBD3 complex

The interaction between dmDcr-1 and its cofactor Loqs is mediated by the dsRBD3 which is only present in isoforms Loqs-PA and –PB (Förstemann et al., 2005). When not associated with his dsRBP partner, dmDcr-1 has some weak processing activity for pre-miRNA substrates which is drastically increased upon binding to the isoform Loqs-PB (Fukunaga et al., 2012; Jiang et al., 2005). Additionally, it has been shown that a 46 aa extension of the Loqs-PB specific interdomain linker region could enhance affinity of Loqs to dmDcr-1 (Ye et al., 2007). In this study, the molecular details of the interaction between the dsRBD3 and the Loqs-PB specific interdomain linker to dm-Dcr-1 were investigated. By truncation of the linker region between dsRBD2 and dsRBD3, a stretch of 24 aa together with the dsRBD3 was identified, which contains all residues that are required for the interaction with dmDcr-1 (Figure 2.16). Using different technical approaches, the interaction surface with dmDcr-1 was identified to be separated into two different binding modules. One module is the dsRBD3 itself, which has a rather low affinity to dmDcr-1. The solvent exposed F419 in the loop region L2 from the dsRBD3 thereby plays an essential role (Figure 2.16 A). A point mutation at this positon strongly affects the interaction with dmDcr-1 and the activity in the *in vivo* reconstitution assay.

Another module is located within the dsRBD2/3 interdomain linker within a putative amphipathic helical folding, where one side is made up by several unpolar Leu residues and the other side contains mainly charged residues (Figure 2.17). Several point mutations within this region were identified to have an impact on dmDcr-1 binding and pre-miRNA processing activity of the complex. Altogether these residues make up a strong contribution to the interaction surface which leads to a tenfold increase of the affinity between the Loqs and dmDcr-1 when the Loqs-PB specific linker region is present (Figure 2.18 B). This increased affinity is furthermore required for an effective complex formation and functionality *in vivo*, as only isoform Loqs-PB and not -PA led to a rescue of the mature let-7a miRNA in the rescue assay (Figure 2.21 B).

Taken these findings together, two specific dmDcr-1 contact sites located in the linker region as well as in the dsRBD3 were identified. While the low affinity interaction of the dsRBD3 might be important for specificity, the linker region strongly contributes to a high binding affinity. This twostep model for dmDcr-1 binding gives further rise to the speculation, that a proof reading mechanism might take place during pre-miRNA positioning for an accurate processing and/or strand selection upon Ago1 loading. Such a model, however, needs further experimental verification, for example by solving the crystal-structures of either Loqs-3 alone or the dmDcr-1 helicase surface in complex with Loqs-3. These structural data would provide further insights into cooperative pre-miRNA processing and the involvement of the interdomain-linker region.

For the human Loqs homolog TRBP, there exist structural data about the TRBP/Dicer interface (Wilson et al., 2015). The TRBP construct contains most of the dsRBD2/3 interdomain linker and

a helical insertion of the Dicer helicase domain. The linker region is rather flexible and almost completely not resolved. However there exists a small helical part that is packed against the dsRBD3, which is supposed to originate from the dsRBD2/3 interdomain linker region. It is therefore tempting to speculate that there might also be an additional binding site where this part of TRBP contacts Dicer, which is not present within the Dicer construct that was used by this group for crystallization.

To identify the potential interaction surface of the dmDcr-1 site, a limited proteolysis of dmDcr-1 Loqs-3 complex coupled with MS analysis could be performed. Having identified such a minimal interaction construct, crosslinking of the respective dicer helicase fragment with Loqs-3 could be repeated as crosslinking with the full length dmDcr-1 with Loqs-3 did not result in reliable inter-crosslinks which might be caused by the predominant amount of peptides from dmDcr-1 intra-crosslinks. From the full length dmDcr-1-Loqs-3 crosslink, one peptide could be identified with a relatively weak score, where K391 is crosslinked to K673 of the dmDcr-1 site. These two Lys were located within the interdomain region between the helicase domain and the dsRBD on the dmDcr-1 surface and in front of the dsRBD3 in the Loqs-3 construct. However these observations require further experimental replicates and optimization.

2.3 Dimerization of Loqs-dsRBD3

After solving the crystal structure of Loqs-dsRBD3, the asymmetric unit cell of the crystal revealed an unexpected assembly of the dsRBD3 monomers. Two monomers exchange secondary structure elements and were physically coupled to each other through an exchange of secondary structure elements. The overall folding of the single domain swapped dsRBD3s, however, is identical to the folding observed for each of the three unswapped monomers, which are additionally present in the asymmetric unit cell. The domain swap occurred in loop region L3 and generates two monomers with the topologies α1β1β2β3'α2' and α1'β1'β2'β3α2 (Figure 2.13 A and B). To each domain swapped monomer a second native monomer is assembled generating a symmetric dimer. As the domain swap might rather be an artefact due to lacking crystal contacts that are required for the formation of the macroscopic crystal, the dimeric assembly, however, was identified to be highly stable in solution according to *in silico* calculations with the PISA-server (Krissinel and Henrick, 2007). Consistent with the observations from the data, the purified Loqs-dsRBD3 eluted at a volume that corresponds to twice as much as the molecular weight calculated for one single monomer and could also be readily detected by chemical crosslinking of two Loqs monomers with DSS (Figure 2.22, B and C). A mutation of L426 to Arg, which introduces a charged and bulky sidechain into the mostly hydrophobic surface for dimerization, was able to completely abolish dimer formation *in vitro* and *in vivo* (Figure 2.23, A and B).

The crystallized Loqs-dsRBD3 dimer uses a high confidence surface for the interaction of both monomers when dmDcr-1 is not present for binding. Interestingly, this surface is conserved in several other protein-interacting dsRBDs among different species (Figure 3.2, A). And in fact for the dsRBD3 of TRBP and PACT it has been shown by several other groups that there exists a tendency for the formation of homo- and heterodimers (Hitti et al., 2004; Kok et al., 2007; MacRae et al., 2008). A sequence alignment of the dsRBD3 from dsRBPs acting in small RNA processing pathways in various organisms shows a strong conservation of structural elements that were required for the formation of the Loqs-dsRBD3 dimers (Figure 3.2, A). It is therefore likely that dimerization might either be required for stabilization of a Dicer unassociated dsRBP and/or for any yet unknown physiological processes in the cell which require a Dicer unassociated form of the dsRBP (See working model in Figure 3.3).

The Drosha dsRBP partner DGCR8 from the microprocessor complex has been shown to dimerize and act together with Drosha as a Dimer (Faller et al., 2007; Herbert et al., 2016; Senturia et al., 2012, 2010). To determine the stoichiometry of the dmDcr-1 Loqs-PB complex, single molecule analysis was applied. Bleaching steps of Loqs-3^{AF647} bound to dmDcr-1 were counted and revealed a clear 1:1 ration between dmDcr-1 and Loqs-3 (Figure 2.24). Consistent with this finding, the interaction surface between the dsRBD3 of TRBP and a helical insertion of the Dicer helicase domain also reveals a 1:1 complex of both proteins (Wilson et al., 2015) indicating a dissociation of the dsRBP Dimer upon binding to its respective Dicer partner. A superposition of Loqs-dsRBD3 with TRBP dsRBD3 shows an almost identical folding of both proteins and the single dsRBDs can easily be aligned on top of each other in both structural models (Figure 3.2, B; C and D). Thus it is conceivable that TRBP and PACT might form homo- and/or heterodimers in a similar manner then Loqs and that the interaction surface between Loqs and dmDcr-1 might comprise an analogy to the one from TRBP in complex with the human Dicer.

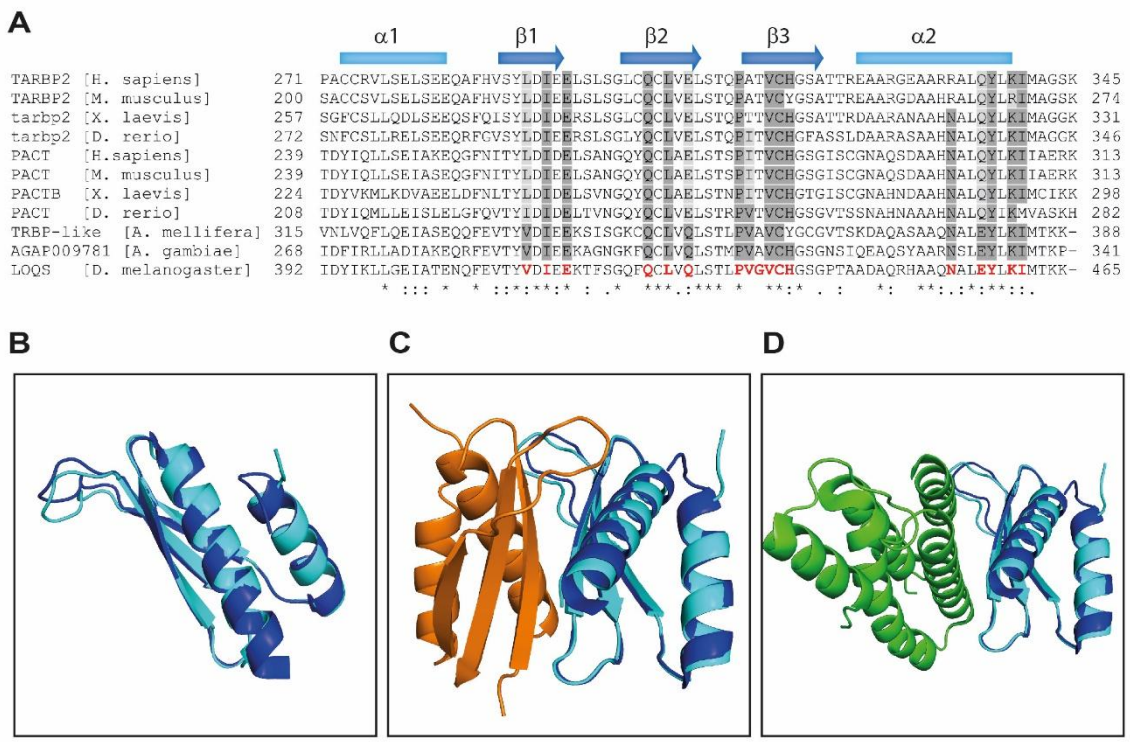


Figure 3.2: The dimerization interface of Loqs-dsRBD3. (A) Sequence alignment Loqs-dsRBD3 with functional homologs of the TRBP and PACT protein families from different organisms. **(B)** Superposition of Loqs-dsRBD3 (cyan) and TRBP dsRBD3 (blue, PDB ID: 4WYQ) **(C)** Superposed structure of Loqs/TRBP dsRBD3 packed against a second Loqs monomer (orange). **(D)** Superposed structure of Loqs/TRBP dsRBD3 packed the human Dicer helicase interface.

Furthermore it has been shown by other groups (Fukunaga et al., 2012; Ye et al., 2007) that the affinity of the dmDcr-1 interacting isoforms from Loqs, Loqs-PA and –PB, strongly differs between these variants. Both isoforms were generated by alternative splicing and Loqs-PB contains an extended interdomain linker region. In this thesis, it has been shown by FCCS measurements that this region is able to increase the affinity around 10 fold compared to the dsRBD3 alone (Figure 2.18, B and C). In the *in vivo* dicer reconstitution assay, the rescue of let-7a miRNA was inefficient when dmDcr-1 was co-transfected with Loqs-PA. Consistent with that, a recent study identified Loqs-PB as the predominant isoform that interacts with dmDcr-1, however which is only required for ~40% of all endogenous miRNAs which are expressed in flies (Lim et al., 2016).

Considering the dimer formation of Loqs-dsRBD3, it is therefore possible that an alternative function of the Loqs dimer might exist *in vivo* when it is not associated with dmDcr-1 (See working model in Figure 3.3). The mRNA expression of *loqs-RA* and *loqs-RB* strongly differs among tissues and developmental stages in *Drosophila* somatic cells (Fukunaga et al., 2012). The expression of *loqs-RA* is in general lower at earlier stages of development and increases with progressing time of development. The germ cells from male and female flies also comprise different expression of both isoforms. In adult male flies, testes predominantly express *loqs-RA*,

whereas female ovaries express *loqs-RB* in higher levels then *loqs-RA* (Fukunaga et al., 2012). If these differences in the expression of the two isoforms originate from a need of different miRNA isoforms during developmental processes or from potential Dicer unrelated functions in the cell is unclear. For the Loqs homolog TRBP it has been shown that besides its involvement in the generation of cellular miRNAs, it has a function in cell cycle regulation, where TRBP undergoes hyperphosphorylation during mitosis by the c-Jun N-terminal kinase (JNK), which then enhances the inhibitory activity of TRBP on PKR (Kim et al., 2014).

If there are further Loqs-mediated functions besides its implication in the miRNA processing pathway, needs to be experimentally verified. An IP of either endogenous or overexpressed Loqs-PA and -PB in different tissues or cell lines from *Drosophila* combined with MS analysis might provide information about other protein interactors and a potential contribution of the Loqs dimer in other cellular processes. Consistent with this assumption, Loqs has been identified in a MS coupled GST pulldown of Sex-lethal (SXL), which is responsible for the female-specific repression of *male-specific-lethal-2* (*msl2*) mRNA (Graindorge et al., 2013). However if the findings of this study are specific and Loqs actually comprises a function within the *msl2* repression needs to be further experimentally tested.

If there exist a free Loqs dimer with Dicer-independent functions, Loqs-PA might be stronger implicated within these processes because of its lower affinity to dmDcr-1 (Figure 3.3 upper part, bold arrow) and Loqs-PB on the other side is rather responsible for the generation of miRNAs in flies (Figure 3.3 middle part, bold arrow). Figure 3.3 summarizes a hypothetical working model of both Loqs isoforms.

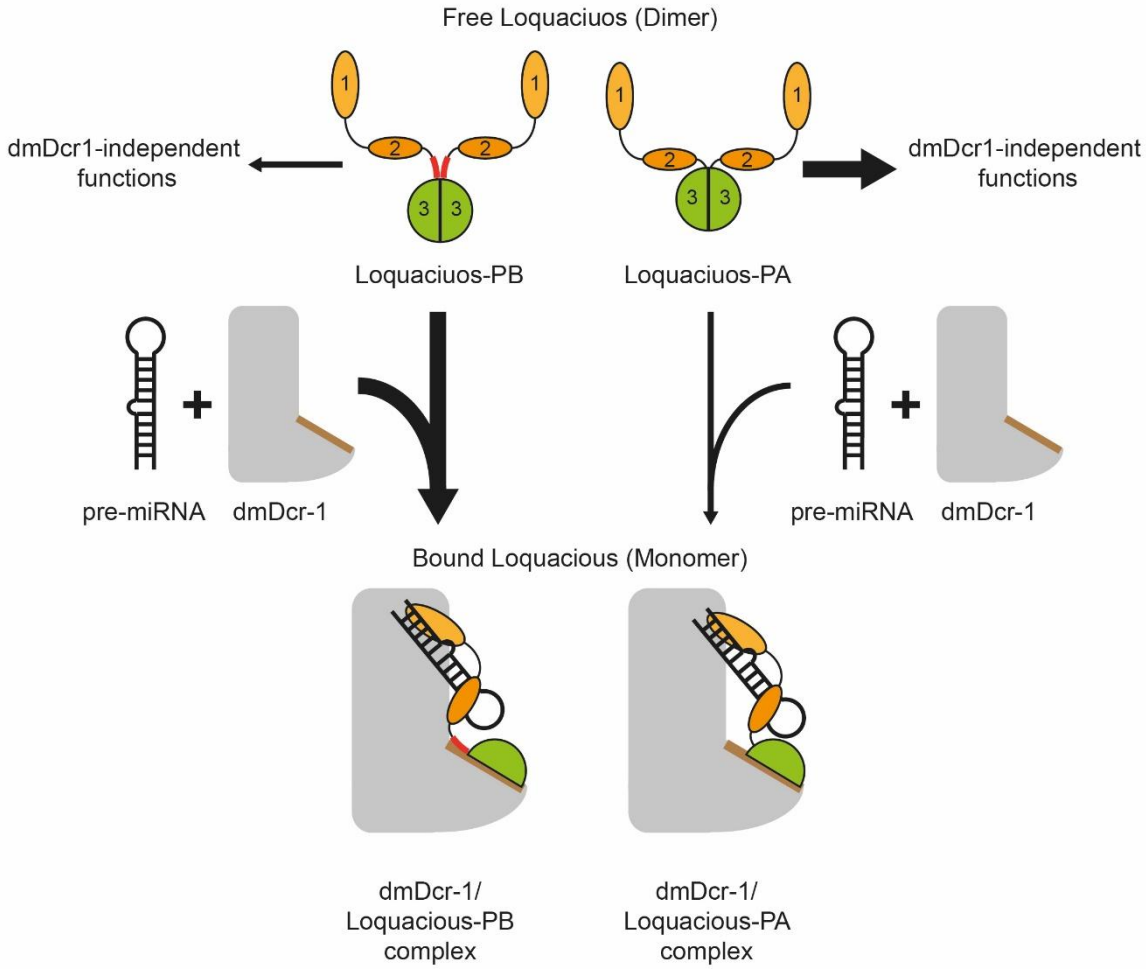


Figure 3.3: Working model for Loqs dimerization and dmDcr1 binding. In the absence of dmDcr1, Loqs forms homodimers via the dimerization interface located on dsRBD3 (green) and might be involved in dmDcr-1 independent processes. DsRBD1 and 2 are indicated as orange circles. In the presence of dmDcr1, Loqs-PB preferentially binds to the helicase domain of dmDcr1 (left, indicated with bold arrow) via the dimerization interface of dsRBD3 and a high-affinity binding site located in the linker preceding dsRBD3 (highlighted in red) leaving Loqs-PA to exercise putative dmDcr-1 independent processes (right, indicated with bold arrow).

2.4 Post-translational modifications of human Dicer

GST-pulldown of the human Dicer with the dsRBD3 of TRBP

For the phosphorylation analysis of human Dicer, a pulldown with a C-terminal part of TRBP was used. The construct contained the dsRBD3 together with a fragment of the dsRBD2/3 interdomain linker (residues 261-364), which was indispensable for the solubility of the protein. A complete truncation of the interdomain linker led to aggregation of the dsRBD3 which further localized into inclusion bodies upon expression in *E. coli*. Therefore, aas from position 261-292 from TRBP have a substantial influence on the stability of the dsRBD3 and presumably of the full length protein. The molecular details behind this stability effect however require further experimental validation. To characterize the structural details of the dsRBD3 and the interdomain linker from TRBP, a crystallization approach was performed within this study. Therefore, TRBP dsRBD3 together with the interdomain linker from position 261 were used and crystal screening and refinement resulted in macromolecular hexagonal protein 3D-crystals (Condition 1: 10% PEG 10000, 100 mM Tris pH 7.8-8.7; Condition 2: 18 % PEG 20000, 100 mM Tris pH 7.8-8.7). These crystals diffracted only to ~3.5 Å resolution and no symmetry could be assigned, which ruled out a solution of the crystal structure. In further experiments, the crystallization conditions could be varied to obtain another crystal, which might be suitable for X-ray crystallography. Additionally nuclear magnetic resonance (NMR) measurements could be performed to characterize the structure of the interdomain linker in solution and its contribution to the solubility of the protein. Consistently with that, the published crystal structure of the Dicer/TRBP interface (Wilson et al., 2015) contains major parts of the interdomain linker and an unknown electron density which might derive from this interdomain linker located next to one TRBP monomer comprising a helical structure element, that might derive from the putative amphipathic helix that contains several identical residues as in Loqs-PB.

Analysis of post-translational modifications on hDcr

For the precipitation of either the endogenous or a FLAG/HA tagged overexpressed hDcr, a GST-tagged TRBP truncation (261-364), containing the 3rd dsRBD and a part of the dsRBD2/3 interdomain linker was used. The GST-pulldown of hDcr was suitable to obtain sufficient sample for MS analysis, which could not be obtained by Dicer-IP or other strategies used for the enrichment of hDcr so far. The MS data from one single measurement resulted in several phospho-sites for the overexpressed FLAG/HA hDcr with reliable scores and one single site for the endogenous hDcr reflecting the amounts of precipitated protein that was used for the measurements and which was higher in the case of the overexpressed hDcr. To further validate the identified phospho-sites, more replicates of the pulldown experiment are required to confirm

the biological relevance of these sites. Additionally for endogenous hDcr, the pulldown needs to be further optimized, to obtain more material for the identification of additional endogenous phospho-sites.

All identified phospho-sites were located within or in close proximity of domains from the hsDcr, which were involved in the binding and processing of dsRNA. Thereby, Y153 is located in the helicase domain, S678 in the dsRBD, S1016 in the PAZ domain, Y1330 and S1470 in the RNase III domain. The S1252 and the adjacent S1255 are located in the interdomain region between the PAZ and RNase III domain und thus might also have an effect on the hsDcr function. For the characterization of the identified phospho-sites, a similar rescue assay, which was used for the analysis of different Loqs-PB mutants, could be applied. The mutations on the hsDcr sequence could be introduced by site directed mutagenesis and the putative phospho-site should be mutated to Glu, which mimics a phosphorylation at the respective position, or to Ala which cannot be phosphorylated by cellular kinases and mimics the unphosphorylated state of the protein. The amount of rescued miRNAs in MEF $Dcr^{-/-}$ cells would provide evidence on the regulatory effects of a certain mutation on the processing activity of hsDcr. Especially phospho-sites Y153 and S678 have not been annotated yet on the PhosphoSitePlus® (www.phosphosite.org) databank from any screens for post-translational modification and therefore might be some interesting candidates for further validation. However, also from the annotated phospho-sites, most data is derived from more globular phospho-proteomic screens and no data based on site-specific methods, e.g. site directed-mutagenesis, are available to date.

Furthermore it has been shown that the GST-tagged TRBP truncation (261-391) is not only able to precipitate the hsDcr, but also has a quite strong affinity to the fly homolog dmDcr-1 (Figure 2.25, B). Therefore the dsRBD3 construct of TRBP can also be used to precipitate other Dicer homologs from different cellular lysates in a similar manner as described in Hauptmann et al., (2015) for a TNRC6 fragment that is able to precipitate several Ago homologs in different organisms. In an MS analysis, the different phospho-sites then could be directly compared to obtain information about the conservation of each site. Additionally, lysates from different tissues can be analyzed to obtain information about the phosphorylation pattern within these tissues.

3. The crystal structure of Brat-NHL in complex with its consensus RNA motif

In post-transcriptional gene regulation processes, many RBPs interact with target RNA sequences. Thereby the specificity and selectivity of this interaction is of high relevance for the correct biological function in the cell. To elucidate the molecular interactions between the NHL domain of Brat and its RNA consensus sequence, the crystal structure of the RNA bound complex has been solved at 2.3 Å resolution. The six-bladed β propeller of the NHL domain has a diameter of ~47 Å and a modular shape that provides a rigid binding platform for the interaction with its target RNA. The binding cavity runs along the complete positively charged top surface and forms three binding pockets to specifically interact with the six bases of the consensus motive. This is achieved by base-specific hydrogen bonds to the protein main chain and several side chains within the NHL domain. By the arrangement of the six-bladed β propeller, bases 1 and 2 of the consensus motif were coordinated between blades II and III, base 3 between blades IV and V and bases 4-6 between blades I and VI. Thus, it is tempting to speculate, that the arrangement of the β blades and their loop- and sidechain-composition within the propeller might be crucial for the generation of sequence specificity and substrate recognition.

The β propeller folding of the Brat NHL domain is also a common structural feature in WD40 domains, which are highly abundant in the eukaryotic proteome (Stirnimann et al., 2010). Based on our findings it is tempting to speculate that WD50 might also be a widespread RNA-binding domain. In the past, this domain was mainly regarded as a mediator of protein-protein contacts and only few candidates were shown to interact directly with RNA substrates as the WD40 domain of Gemin5 that interacts with snRNAs (Lau et al., 2009). From recent large-scale proteomic screens, the number of new candidates containing a WD40 domain that directly contacts RNA could be increased (Baltz et al., 2012; Castello et al., 2012; Gerstberger et al., 2014; Kwon et al., 2013). Based on the solved structure of the NHL domain of Brat in complex with its consensus motif, structural homology modeling could be performed to obtain information how sequence specificity is achieved in other WD40 or NHL-domain containing proteins.

Additionally to the structure of the NHL domain of Brat bound to its consensus motive, there exits another monomer which is bound to a stretch of poly U that directly succeeds the consensus motive. The poly U stretch adapts a similar conformation as the consensus motif and can undergo most of the respective interactions. However uracil is unable to form any of the hydrogen bonds, observed for the guanine base, which affects mostly the binding of U14 that is equivalent with G6 from the consensus motif, as the electron density for this nucleotide is only very weakly defined. The orientation of the two Brat NHL monomers within the crystal seems rather influenced by the simultaneously bound RNA, as the little protein contacts within the crystal are not indicative for a

stable interaction between both monomers. However, as the full length Brat contains a CC, which is used for dimerization by proteins from the TRIM family, it is not excluded that this orientation of both NHL domains has function in the Brat-mediated translational repression *in vivo*. This possibility could be addressed in further experiments with the full length protein.

IV. Material and Methods

1. Material

1.1 Chemicals and enzymes

If not denoted separately, chemicals were ordered from Amersham Biosciences (Buckinghamshire, UK), Applichem (Darmstadt, Germany), Biorad (Hercules, USA), Merck (Darmstadt, Germany), Roth (Karlsruhe, Germany) and Sigma Aldrich (Munich, Germany). DNA oligonucleotides were ordered and synthesized from Metabion (Martinsried, Germany), RNA oligos from Biomers (Ulm, Germany) and Eurogentec (Liège, Belgium). All enzymes which were used for cloning and modification of DNA or RNA were purchased from Thermo Fisher Scientific (Waltham, USA) or New England Biolabs (Ipswich, USA).

1.2 Laboratory equipment

All instruments which were used in this were listed in the following table.

Table 4.1: Instruments

Instrument	Distributor (alphabetical)
Ultraspec 3300 pro	Amersham Biosciences (Little Chalfont, UK)
GeneAmp PCR System 9700	Applied Biosystems (Foster City, USA)
Avanti J-20 XP Centrifuge	Beckman Coulter (Krefeld, Germany)
Geiger Counter LB123 EG&G	Berthold (Bad Wildbad, Germany)
CFX96 Real-Time-System	Bio-Rad (Hercules, USA)
PowerPac HC Power Supply	Bio-Rad (Hercules, USA)
Screen Eraser-K	Bio-Rad (Hercules, USA)
Trans-Blot SD	Bio-Rad (Hercules, USA)
Power Supply EV233	Consort (Turnhout , Belgium)
Thermomixer compact	Eppendorf (Hamburg, Germany)
Centrifuge 5415D	Eppendorf (Hamburg, Germany)
Polymax 2040	Heidolph (Schwabach, Germany)
Vortexer REAX top	Heidolph (Schwabach, Germany)
Branson Sonifier 450	Heinemann (Schwäbisch Gmünd, Germany)
Biofuge pico	Heraeus (Hanau, Germany)
Hybridization oven T 5042	Heraeus (Hanau, Germany)
Incubator Model B6200	Heraeus (Hanau, Germany)
Megafuge 40	Heraeus (Hanau, Germany)
Centrikon T-1170	Kontron
Microscope Diavert	Leica (Wetzlar, Germany)
AT200	Mettler (Gießen, Germany)
Milli-Q PLUS	Millipore (Billerica, USA)
Innova 44	New Brunswick Scientific (Eppendorf)
Quantum gel-doc. system	Peqlab (Erlangen, Germany)
LC 4801 P	Satorius (Göttingen, Germany)
Cytoperm	Thermo Scientific (Rockford, USA)

HeraCell 240i CO2 Incubator	Thermo Scientific (Rockford, USA)
Nanodrop	Thermo Scientific (Rockford, USA)

1.3 Plasmids

The following vectors were used in this thesis.

Table 4.2: Plasmids for bacterial expression

Plasmid	Resistance	Description
pGEX4T-1	AMP	N-terminal GST
pET32a	AMP	N-terminal His_6-Trx
pColdI	AMP	N-terminal His_6, Cold shock inducible
pETDuet-1	AMP	His-(MCS1) and S-tag(MCS2), co-expression of two target genes
pHis.*TEV*	AMP	Expression vector for His_6 TEV-protease
pGEX.*TEV*	AMP	Expression vector for GST TEV-protease

Table 4.3: Plasmids for eukaryotic expression

Plasmid	Resistance	Description
VP5-FAME	AMP,Neo	Modified pIRES neo, N-terminal FLAG and HA, *FseI-AscI* cassette added
pCS2-FAME	AMP	N-terminal myc_6-tag, *FseI-AscI* cassette added
pGTO-b	AMP	Modified pCDNA3/TO, eGFP inserted for the generation of stable HEK T-REx-293 cell lines

Table 4.4: Plasmids for insect expression

Plasmid	Resistance	Description
pMT/V5-A	AMP	Copper inducible expression vector for *Drosophila* derived cells
pFastBac-1	AMP	Expression in SF21 cells with Bac-to-Bac® expression system

1.4 Oligonucleotides

All DNA- or RNA-oligonucleotides that were used for this thesis are listed in the following tables.

1.4.1 DNA oligonucleotides

All DNA oligonucleotides, which were used in this thesis, were ordered from Metabion international AG (Munich).

Table 4.5: List of different DNA oligos.

Primer name	Sequence 5'->3'
Loquacious	
Loqs-PB_338-463for	GCTAGGATCCGAAAACCTGTATTTTCAGGGACCCCGCAGTAGTGAAAATTATTATGG
Loqs-PB_348-463for	GCTAGGATCCGAAAACCTGTATTTTCAGGGATTGAAAGATATCTCTGTGCCG

Loqs-PB_358-463for	GCTAGGATCCGAAAACCTGTATTTTCAGGGAACGCAGCACAGTAACAAAGTATCC
Loqs-PB_368-463for	GCTAGGATCCGAAAACCTGTATTTTCAGGGACATAAGACCCTAAAAAATGC
Loqs-PB_379-463for	GCTAGGATCCGAAAACCTGTATTTTCAGGGACTGCTTAAGTTACAGAAGACTTGC
L-PBf Gly2 Linker for	CCAGGATCCGAAAACCTGTATTTTCAGGGAGGAGGAATTGATTACATCAAGCTGCTGG
Loqs-PB_392-463 rev	CCAGTCGACTTAGGTCATGATCTTCAAGTACTCG
L-PB-463-HA rev	CCAGTCGACTTAAGCGTAATCTGGAACATCGTATGGGTAGGTCATGATCTTCAAGTACTCG
L-PBLink338-391rev	CTAGGTCGACTTACTTGTTGTTCTTCAAGCAAGTC
L-PB_k FseI for	GCTAGGCCGGCCATGGACCAGGAGAATTTCC
Loqs-PB_392-463 rev	TAGCGGCGCGCCTTAGGTCATGATCTTCAAGTACTCG

Dicer

GFP_pFB_f	GCTAGAATTCACCATGCATCACCATCACCATCACGGCGGCAGCAAGGGCGAGGAGCTGTTCACC
GFP_pFB_EcoRI_r	CTAGGAATTCCTTGTACAGCTCGTCCATGCC
Dicer1_FseI for	GCTAGGCCGGCCATGGCGTTCCACTGGTGCG
Dicer1_AscI rev	TAGCGGCGCGCCTTAGTCTTTTTTGGCTATCAAGCC

TRBP

TRBP_231-End for	GATCGGATCCGAAAACCTGTATTTTCAGGGAGATGCCCGGGATGGCAATGAGG
TRBP_241-End for	GATCGGATCCGAAAACCTGTATTTTCAGGGAGATGATGACCACTTCTCCATTGG
TRBP_251-End for:	GATCGGATCCGAAAACCTGTATTTTCAGGGATCCCGCCTGGATGGTCTTCG
TRBP_261-Gly2 for	GATCGGATCCGAAAACCTGTATTTTCAGGGAGGAGGACCAGGTTGCACCTGGGATTCTCTACG
TRBP_271-End for	GATCGGATCCGAAAACCTGTATTTTCAGGGATCAGTAGGAGAGAAGATCCTGTCC
TRBP_231-291 rev	GATCGTCGACTTAGCCCAGGGCACCCAGGGAGCC
TRBP_292-364 rev	CCAGTCGACTTAGCCTGCCATGATCTTGAGGTACTGC
TRBP_k FseIfor	GCTAGGCCGGCCATGCTGGCCGCCAACCCAGGC
TRBP_k AscIrev	TAGCGGCGCGCCTCACTTGCTGCCTGCCATGATCTTGAGG

Ago

Ago1 (EcoRI) for	CCAGAATTCATGGAAGCGGGACCCTCG
Ago1 (HindIII) rev	CCAAAGCTTTCAAGCGAAGTACATGGTGC
Ago2 (EcoRI) for	ATCTGAATTCATGTACTCGGGAGCCGGCCCCGC
Ago2 (HindIII) rev	ATCTAAGCTTTTATCAAGCAAAGTACATGGTGC
Ago3 (EcoRI) for	CCAGAATTCATGGAAATCGGCTCCGCAGG
Ago3 (HindIII) rev	CCAAAGCTTTTAAGCGAAGTACATTGTGC
Ago4 (EcoRI) for	CCAGAATTCATGGAGGCGCTGGGACCC
Ago4 (ClaI) rev	CCAATCGATTCAGGCAAAATACATCGTGTGC

Mutagenesis

L-PB F419A f	CATAGAGGAGAAGACCGCGTCTGGCCAGTTCCAG
L-PB F419A r	CTGGAACTGGCCAGACGCGGTCTTCTCCTCTATG

L-PB F423A f	GAAGACCTTCTCTGGCCAGGCGCAGTGCCTGGTTCAACTG
L-PB F423A r	CAGTTGAACCAGGCACTGCGCCTGGCCAGAGAAGGTCTTC
L-PB A444K f	CAGCGGACCAACAAAGGCCGATGCCCAGC
L-PB A444K r	GCTGGGCATCGGCCTTTGTTGGTCCGCTG
L-PB L371A f	CAGGGACATAAGACCGCAAAAAATGCAACGGGC
L-PB L371A r	GCCCGTTGCATTTTTTGCGGTCTTATGTCCCTG
L-PB N373E_for	GGACATAAGACCCTAAAAGAAGCAACGGGCAAAAAACTG
L-PB N373E_rev	CAGTTTTTTGCCCGTTGCTTCTTTTAGGGTCTTATGTCC
L-PB K378E_for	CTAAAAAATGCAACGGGCAAAGAACTGCTTAAGTTACAGAAGAC
L-PB K378E_rev	GTCTTCTGTAACTTAAGCAGTTCTTTGCCCGTTGCATTTTTTAG
L-PB L379/380/382A f	CAACGGGCAAAAAAGCAGCTAAGGCACAGAAGACTTGCT
L-PB L379/380/382A r	AGCAAGTCTTCTGTGCCTTAGCTGCTTTTTTGCCCGTTG
L-PB L387A f	TAAGTTACAGAAGACTTGCGCCAAGAACAACAAGATTGATTAC
L-PB L387A r	GTAATCAATCTTGTTGTTCTTGGCGCAAGTCTTCTGTAACTTA
L-PB KI460_61EA_FAf	GAATGCCCTCGAGTACTTGGAGGCCATGACCTAAGGCGC
L-PB KI460_61EA_FAr	GCGCCTTAGGTCATGGCCTCCAAGTACTCGAGGGCATTC
L-PB KI460_61EA_SalIf	CAGAATGCCCTCGAGTACTTGGAGGCCATGACCTAAGTCGAC
L-PB KI460_61EA_SalIr	GTCGACTTAGGTCATGGCCTCCAAGTACTCGAGGGCATTCTG
L-PB K460E_FAf	GAATGCCCTCGAGTACTTGGAGATCATGACCTAAGGCGC
L-PB K460E_FAr	GCGCCTTAGGTCATGATCTCCAAGTACTCGAGGGCATTC
L-PB K460E_SalIf	CAGAATGCCCTCGAGTACTTGGAGATCATGACCTAAGTCGAC
L-PB K460E_SalIr	GTCGACTTAGGTCATGATCTCCAAGTACTCGAGGGCATTCTG
L-PB C386S_for	CTTAAGTTACAGAAGACTTCTTTGAAGAACAACAAGATTG
L-PB C386S _rev	CAATCTTGTTGTTCTTCAAAGAAGTCTTCTGTAACTTAAG
L-PB C425S_for	GGCCAGTTCCAGAGCCTGGTTCAAC
L-PB C425S_rev	GTTGAACCAGGCTCTGGAACTGGCC
L-PB C437S_for	CCCGTTGGCGTTAGCCACGGCAGCG
L-PB C437S_rev	CGCTGCCGTGGCTAACGCCAACGGG
TRBP251_-Bam f	CCTGTATTTTCAGGGTTCCCGCCTGGATGG
TRBP251_-Bam r	CCATCCAGGCGGGAACCCTGAAAATACAGG
Ago2 S798A for	GCACACGCTCCGTGGCCATCCCAGCGCCAGC
Ago2 S798A rev	GCTGGCGCTGGGATGGCCACGGAGCGTGTGC

Northern probes

let-7a	AACTATACAACCTACTACCTCA

Q-PCR oligos

pRT GAPDH fwd	TGGTATCGTGGAAGGACTCATGAC
pRT GAPDH rev	ATGCCAGTGAGCTTCCCGTTCAGC
pRT PLK1 Primer1 fwd	AGCAAGAAAGGGCACAGTTT
pRT PLK1 Primer1 rev	GGTTTGCCCACTAACAAGGT

pRT PLK1 Primer2 fwd	ACAGTGTCAATGCCTCCAAG
pRT PLK1 Primer2 rev	CTGATACCCAAGGCCGTACT

1.4.2 RNA oligonucleotides

RNA oligonucleotides were ordered from biomers.net GmbH (Ulm, Germany) or synthetized at the Max Planck Institute (Martinsried, Munich).

Table 4.6: : List of different RNA oligos

siRNA	Guide sequence 5'->3'	Passenger sequence 5'->3'
Si428	UCAUAUUCGACUUUGGUUGdCdC-Cy5	CAACCAAAGUCGAAUAUGAdTdT
Si856	CUUGCCCGCGAACACCUCCdTdT-Cy5	GGAGGUGUUCGCGGGCAAGdTdT
Impβ si1	UUUCCCUUAGCUCAUUCAGUdT	CUGAAUGAGCUAAGGGAAAUdT
Impβ si2	UUACUAGGCUCUUCUUUCCUdT	GGAAAGAAGAGCCUAGUAAUdT
Imp11 si1	UAAUUCUAAACCAUCUUCCUdT	GGAAGAUGGUUUAGAAUUAUdT
Imp11 si1	UAUCUCUACACACAUUAUCUdT	GAUAAUGUGUGUAGAGAUAUdT
Tnpo1 si1	UAAGGCUGCAUCUCUAUACUdT	GUAUAGAGAUGCAGCCUUAUdT
Tnpo1 si2	AUAAAUGAAUACCUGUGGCUdT	GCCACAGGUAUUCAUUUAUUdT
RanBP1 si1	UUGGUCUCCUCCUUCACCGUdT	CGGUGAAGGAGGAGACCAAUdT
RanBP1 si2	UUGUUUUGAAUUUCUGUGCUdT	GCACAGAAAUUCAAAACAAUdT

1.5 Antibodies

All commonly used antibodies were diluted in TBS-T supplemented with 5 % milk powder (Roth) and 0.02% Sodium azide (Sigma Aldrich).

Table 4.7: Antibodies

Antibody	Description	Dilution
Primary antibodies:		
Mouse anti-αTubulin	monocl, clone DM1A, Sigma Aldrich	1:10000
Mouse anti-βActin	monocl, clone AC15, Abcam	1:10000
Mouse anti-HA	monocl, clone 16B12, Covance	1:1000
Mouse anti-GFP	Monoclonal, clones 7.1 and 13.1, Roche	1:1000
Rabbit anti-c-Myc	polycl, C3956, Sigma Aldrich	1:1000
Rabbit anti-dmDcr-1	Polycl., ab4735, Abcam	1:1000
Rabbit anti-hsDcr	Polycl., A301-937A, Bethyl	1:5000
Rat anti-Ago2	Monoclonal, hybridoma Ab (clone 11A9, Rüdel et al.,2008)	1:10
Secondary antibodies:		
Goat anti-rat 680	secondary antibody, Li-Cor Biosciences	1:10000
Goat anti-rabbit 680	secondary antibody, Li-Cor Biosciences	1:10000
Goat anti-mouse 680	secondary antibody, Li-Cor Biosciences	1:10000
Goat anti-rat 800	secondary antibody, Li-Cor Biosciences	1:10000
Goat anti-rabbit 800	secondary antibody, Li-Cor Biosciences	1:10000
Goat anti-mouse 800	secondary antibody, Li-Cor Biosciences	1:10000

1.6 Bacterial strains and cell lines

Table 4.8: Bacterial strains

Strain	genotype	distributor
BL21 (DE3)	*F^- dcm ompT lon hsdSB(rB^- mB^-) gal λ(DE3)*	Invitrogen
Rosetta (DE3)	*F^- dcm ompT lon hsdSB(rB^- mB^-) gal λ(DE3) pRARE (CamR)*	Novagen
DH10B	*F^- endA1 $deoR^+$ recA1 galE15 galK16 nupG rpsL Δ(lac)X74 φ80lacZΔM15 araD139 Δ(ara,leu)7697 mcrA Δ(mrr-hsdRMS-mcrBC) Str^R λ^-*	Invitrogen
XL1 *blue*	*F^- recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac F'[$proAB^+$ $lacI^q$ ZΔM15 Tn10 (Tet^R)]*	Stratagene

Table 4.9: Eukaryotic cell lines

Cell line	specification
HEK 293T	Human embryonic kidney cells
MEF Dcr $^{-/-}$ and $^{+/+}$	Mouse embryonic fibroblasts
d.mel-2	Insect cells derived from *Drosophila melanogaster*
Sf21	Insect cells derived from *Spodoptera frugiperda*

1.7 Culture media

Standard cell culture medium for human cells:

Dulbecco's Modified Eagle Medium (DMEM) (Sigma Aldrich, Munich, Germany)

Supplemented with:

10 % fetal bovine serum (FBS) (Sigma Aldrich, Munich, Germany)

1 % penicillin/streptomycin (Sigma Aldrich, Munich, Germany)

Standard cell culture medium for insect cells:

Gibco SF900 medium (Thermo Scientific) for Sf21 cells and Express Five® SFM medium (Thermo Scientific) supplemented with GlutaMAX™ (Thermo Scientific) for d.mel-2 cells.

Standard cell culture medium for bacteria cells:

Antibiotics for selection of positively transformed bacteria cells were added from 1000 x stock solutions to the final concentrations of 50 µg/ml Ampicillin (Amp), 30 µg/ml Kanamycin (Kan) and 25 µg/ml Chloramphenicol (Cam).

LB-Medium:

1 % NaCl

0.5 % Yeast-Extract

10 % Bacto-Trypton

LB-Plates:

LB medium supplemented with 1.5 % Agar

1.8 Buffers and solutions

Common buffers and solutions

Coomassie staining-solution	30 % EtOH 10 % Acetic acid 0.25 % Coomassie brilliant blue R250
Coomassie destaining-solution	30 % EtOH 10 % Acetic acid
Easy Prep	10 mM Tris/HCl pH 8.0 1 mM EDTA 15 % (w/v) saccharose 2 mg/mL lysozyme 0.2 mg/mL RNase A 0.1 mg/mL BSA
EDC (1-ethyl-3-(3-dimethyl-aminopropyl)-carbodiimid) crosslinking solution	160 mM EDC 130 mM 1-methylimidazole 12.5 mM HCl
Gradient Buffer	25 mM Tris/HCl pH 7.5 150 mM KCl 2 mM EDTA
HEPES buffered saline (HBS) (2x)	54,6 mM HEPES/NaOH pH 7.1 274 mM NaCl 1,5 mM Na_2HPO_4 x 2 H_2O
IP-lysis	25 mM Tris/HCl pH 7.5 150 mM KCl 2 mM EDTA 0.5 % NP-40
IP-wash	50 mM Tris/HCl pH 7.5 300 mM NaCl 5 mM $MgCl_2$ 0.05 % NP-40
Laemmli loading dye (5x)	300 mM Tris/HCl pH 6.8 10 % SDS 62.5 % Glycerin 0.025 % bromophenol blue 10 % β-mercaptoethanol
Northern blot hybridization	20 mM Na_2HPO_4 pH 7,2 15 mM Trinatriumcitrate-dihydrate 150 mM NaCl 7 % SDS 0,02 % Albumin fraction V 0,02 % Polyvinylpyrrolidon K30 0,02 % Ficoll 400
Northern blot wash I	75 mM Trinatriumcitrate-dihydrate pH 7.0 750 mM NaCl 1 % SDS

Northern blot wash II	15 mM Trinatriumcitrate-dihydrate pH 7.0 150 mM NaCl 1 % SDS
PBS	12 mM Na_2HPO_4/NaH_2PO_4 pH 7.5 137 mM NaCl 2.7 mM KCl
RNA sample buffer (2x)	99.9 % Formamide 0.05 % Xylene cyanol 0.05 % Bromophenol blue
Roeder A	10 mM HEPES/NaOH pH 7,9 10 mM KCl 1,5 mM $MgCl_2$ 0,5 mM DTT 1 mM AEBSF
SDS running buffer	25 mM Tris 192 mM Glycine 0.1 % SDS
siRNA annealing buffer (2x)	60 mM HEPES/NaOH pH 7,4 4 mM MgAc 200 mM KAc
TBE	90 mM Tris pH 8.0 90 mM Boric-acid 2 mM Na_2-EDTA
TBS	10 mM Tris/HCl pH 8.0 150 mM NaCl
TBS-T	10 mM Tris/HCl pH 8.0 150 mM NaCl 1 % Tween
Towbin blotting buffer	25 mM Tris/HCl pH 8.6 192 mM glycine (10 % MeOH for semi-dry blotting)

Buffers for protein purification

GST-A	50 mM Tris/HCl pH 8.0 200 mM NaCl
GST-B	50 mM Tris/HCl pH 8.0 200 mM NaCl 10 mM glutathione
HisA	50 mM Tris/HCl pH 8.0 300 mM NaCl 10 mM imidazole
HisB	50 mM Tris/HCl pH 8.0 300 mM NaCl 300 mM imidazole

SEB1	20 mM Hepes/NaOH pH 7.5 150 mM NaCl 1 mM DTT
SEB2	20 mM Tris/HCl pH 7.5 200 mM NaCl 1 mM DTT
SEB3	20 mM Tris/HCl pH 7.5 1 mM DTT
SEB4	20 mM Tris/HCl pH 8.0 200 mM NaCl

1.9 Columns for protein purification

Columns for affinity chromatography

Column	Material	h [mm]	Ø [mm]	V [mL]
IMAC I	IMAC Sepharose™ 6 Fast Flow (GE)	25	16	5
IMAC II	IMAC Sepharose™ 6 Fast Flow (GE)	200	16	11
IMAC III	Profinity™ IMAC (Biorad)	200	26	45
GSTrap	Glutathione Sepharose™ 4 Fast Flow (GE)	25	16	5
GST II	Glutathione Sepharose™ 4 Fast Flow (GE)	200	16	18

Columns for size exclusion chromatography

Column	Material	M_r [kDa]	Ø [mm]	V_t [mL]
S75 10/300 (GE)	Crosslinked agarose and dextran	3-70	10	24
S75 16/600 (GE)	Crosslinked agarose and dextran	3-70	16	120
S75 26/600 (GE)	Crosslinked agarose and dextran	3-70	26	320
S200 10/300 (GE)	Crosslinked agarose and dextran	10-600	10	24
S200 26/600 (GE)	Crosslinked agarose and dextran	10-600	26	320
Superose6 10/300 (GE)	Composite of crosslinked agarose	5-5000	10	24
HiPrep Desalt 26/60 (GE)	Sephadex G-25 Fine	1-5	26	53

2. Molecular biological methods

2.1 Agarose gel electrophoresis

For the detection, separation and analysis of DNA fragments, agarose gel electrophoresis was performed on 0.6-2% (w/v) agarose gels supplemented with ethidium bromide (EtBr) (3µL / 100 mL agarose gel) for staining of the DNA bands. Prior loading to the gel, 6x DNA loading dye (Fermentas) was added to the DNA sample. The separation was performed in TBE buffer for 20-40 min (RT) with a constant current of 120 V. For size determination of DNA bands, a marker (GeneRuler, Fermentas) was loaded. The DNA bands were visualized with a gel documentation system (Quantum ST4, PeqLab (Erlangen, Germany)) and images were stored locally.

2.2 Preparation, purification and sequencing of plasmid DNA

Plasmid-DNA was extracted from chemically competent *E. coli* XL1 blue cells, which were prepared and transformed as described in (Inoue et al., 1990). The DNA was isolated using either the Nucleobond Plasmid Kit for Mini- or the Nucleobond Xtra Midi Kit (both Macherey-Nagel, Düren, Germany) for Maxi-preparations. The quality and concentration of the DNA was determined using NanoDrop 1000 Spectrophotometer (Thermo Scientific, Rockford, USA). For sequencing, samples were prepared according to companies protocols and sent to GATC Biotech (Cologne, Germany) or Macrogen (Amsterdam, Netherlands).

For screening of lager amounts of clones, a fast DNA extraction method was applied according to (Berghammer and Auer, 1993). Therefore 2 mL an overnight culture were resuspended in 40 µL EasyPrep buffer for high copy vectors or 20 µL for low copy vectors and incubated for 1 min at 99°C. After an incubation of 1 min on ice, the supernatant was cleared by centrifugation at 15000 g for 15 min (RT). 5-10 µL of the plasmid containing supernatant was used for an analytical digestion (see 4.2.1.3).

2.3 Restriction analysis

Analytical digestion

For the analytical restriction analysis the following digestion mix was used:

12 µL 10x reaction buffer

8 µL H_2O

1 µL of respective restriction enzyme(s)

2 µL of digestion mix was added to a total volume of 10 µL DNA sample and incubated at 37°C for 1 h followed by agarose gel electrophoresis. Plasmid DNA from positive clones was isolated via Mini-prep purification and stored at -20°C.

Preparative digestion

For preparation and further ligation of DNA from a PCR fragment or a circular plasmid, the following preparative digestion mix was prepared:

2-5 µg	Template DNA
5 µL	10x reaction buffer
ad 50 µL	ddH_2O

The reaction mix was incubated for 1 h at 37°C and digested DNA was purified from agarose gel using the NucleoSpin Gel and PCR Clean-Up Kit (Macherey-Nagel).

2.4 Molecular cloning

DNA sequences for cloning were amplified via polymerase chain reaction (PCR) either from template vectors or from cDNA of *Drosophila Melanogaster* embryonic cells using the Phusion High-Fidelity DNA Polymerase (Thermo Scientific). The PCR mix and the 2-step reaction were performed as below:

PCR-mix

50 ng	Template DNA
2.5 µL	Fwd. primer (10µM)
2.5 µL	Rev. primer (10µM)
10 µL	5x HF buffer
1.5 µL	DMSO (100%)
1 µL	dNTPs (10mM)
1 µL	Phusion (2 U/µL)
Ad 50µL	dd H_2O

Temperature program

time	temperature	cycles
30 sec	98°C	1x
10 sec 30 sec/1kb	98°C 72°C	35x
10 min store	72°C 4°C	1x

After the PCR the mixture was subjected to agarose gel electrophoresis and bands containing the PCR-product were cut out and purified with the NucleoSpin Gel and PCR Clean-Up Kit and preparative digested together with the respective vector for insertion. To prevent re-ligation of the linearized vectors, the 5' and 3' ends were dephosphorylated using the FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific) which was directly added to the digestion (30 min, 37°C). The digested DNA fragments were purified for a second time with the NucleoSpin Gel and PCR Clean-Up Kit.

For ligations, 50 ng of the template vector were mixed with a 2-5 fold molar excess of insert-DNA. After the addition 0.1 U/µL T4 DNA Ligase and the respective buffer (both Thermo Scientific), the reaction was performed at RT for 1 h. For higher efficiency of the ligation reaction, PEG4000 was additionally added in some minor cases. The reaction mixture was heat-inactivated at 65°C for 10 min and transformed into XL1 *blue* and plated onto agarose plates containing the selection antibiotic(s). Positive clones were identified by an analytical digestion and sequenced after plasmid isolation.

2.5 Mutagenesis PCR

For site directed mutagenesis, primers were generated using a web-based program designed to automate the design of mutagenic PCR (PrimerX). The mutagenesis PCR reaction was performed as below:

PCR-mix

50 ng	Template DNA
10 µL	5x HF buffer
1 µL	Fwd. primer (10µM)
1 µL	Rev. primer (10µM)
1.5 µL	DMSO (100%)
1 µL	dNTPs (10mM)
1 µL	Phusion (2 U/µL)
Ad 50µL	dd H_2O

Temperature program

time	temperature	cycles
30 sec	98°C	1x
10 sec	98°C	20x
20 sec	55°C	
30 sec/1kb	72°C	
10 min	72°C	1x
store	4°C	

After the PCR, the sample was purified with the NucleoSpin Gel and PCR Clean-Up Kit and eluted with 44 µL ddH_2O. The template vector was removed by DpnI digestion. Therefore 1µL of the Enzyme and 5 µL of the 10 FD buffer (both Thermo Scientific) was added to the elution and incubated for 3-24 h at 37°C. 5 µL of this mix was used for transformation in XL1 *blue*. After plating onto agarose plates containing the selection antibiotic(s), positive clones were identified by an analytical digestion and sequenced after plasmid isolation.

3. RNA-based methods

3.1 RNA isolation

For quantitative real time PCRs (qRT-PCR) and Dicer rescue assays, RNA was extracted directly from transfected cells, which were pelleted prior lysis. Therefore an adequate amount of TRIzol® (Thermo Fisher Scientific) was added to the cell pellets, according to the manufacturer's guidelines. Purified RNA pellets were resuspended in nuclease free water (5 min at 65°C while shaking) and stored at -80 °C for further appilactions.

3.2 RNA separation and Northern blot analysis

RNA samples were prepared by adding equal amounts of 2x RNA sample buffer the respective RNA-dilutions. 12 % urea polyacrylamide gels (Urea Gel Systems, National Diagnostics) were prepared according to the manufacturer's instructions. To allow a complete separation of RNA species on the gel, the gels were run for 30 min at 400 V in TBE buffer. Prior loading of the RNA samples, pockets were washed thoroughly. After addition of ~30µL samples to the respective pockets, gels were run at 400 V for about 1.5 hours.

To control the quality of the RNA extraction and the loading of each lane, urea gels were stained with EtBr in TBE buffer for 10 min. After documentation, blotting was performed using three

Whatman papers on top and below an Amersham Hybond-N membrane (GE Healthcare) and the urea gel. All components were briefly soaked in MilliQ water upon building of the blotting-sandwich. Transfer of the RNAs was conducted at 20 V for 30 min. Then, the miRNA 5' ends were crosslinked to the membrane, using an EDC crosslinking at 50 °C for 1 h. Therefore, the membrane was placed on another EDC-soaked Whatman paper with the RNA side facing up and wrapped in cling film together with the paper. After crosslinking, the membrane was gently washed with water and dried. Northern blots were hybridized with respective radiolabeled probes overnight at 50°C. After the incubation, the blots were washed twice with Northern blot wash I buffer and once with Northern blot wash II buffer. Signals were detected by exposure to a screen and scanning with the phospho-imaging system PMI (Biorad).

3.3 qRT-PCR

DNA contaminants after RNA isolation were removed from qRT-PCR samples upon digestion with DNAse I for 30 min at 37 °C. For each sample, 1 µg of total RNA was used. Then, DNase I was inactivated by EDTA-addition and 10 min incubation at 65 °C. The cDNA was synthesized according to protocols from the First Strand cDNA Synthesis Kit (Thermo Scientific).

qRT-PCR was performed using Sso Fast Eva Green Mix (Biorad) 0,4 µM forward and 0,4 µM reverse primer and cDNA from 50 ng RNA as template. qRT-PCR were run on a CFX96 cycler (Bio-Rad, Hercules, USA) using standard program as given in the SsoFast EvaGreen SuperMix manual with denaturation and annealing/extension times of 5 sec, 40 cycles and a 65-95 °C melt curve.

qRT-PCR data were assessed, using the ΔΔCt method (Livak and Schmittgen, 2001) with GAPDH as reference mRNA. Error bars were calculated from respective technical and biological replicates.

4. Protein-biochemical methods

4.1 Determination of protein concentration

For the determination of protein concentrations in solutions containing a mixture of several different proteins (e.g. cellular lysates), a Bradford Protein Assay (Biorad) was used and performed according to the protocol of the distributor and as described in (Bradford, 1976).

The concentration of a homogenous protein solutions (e.g. after protein purification) was determined by the measurement of the total A_{280} using the NanoDrop ND-1000 UV-Vis spectrophotometer (Peqlab). The extinction coefficient regarding the absorption from Tryptophan, Tyrosine and non-reduced Cysteine and the molecular weight were calculated *in silico* using the

program Protparam (www.expasy.org/tools/protparam.html). The final protein concentration could be obtained with equation 4.1.

$$\beta\left[\frac{mg}{mL}\right] = \frac{A_{280} \cdot f}{d} \quad and \quad f = \frac{M_r}{\varepsilon_{280}} \quad \text{(Equation 4.1)}$$

β	Mass concentration
A_{280}	Absorption at 280 nm
M_r	Relative molecular weight
d	Thickness of cuvette
ε_{280}	Molecular absorption coefficient

Protein specific factors of proteins, which were recombinantly purified in this thesis, were calculated with Protparam (http://web.expasy.org/protparam) and listed in Table 4.10:

Table 4.10: Protein specific factors

Protein	M_r [kDa]	$\varepsilon_{280}[M^{-1}cm^{-1}]$	f
Loqs dsRBD3	8.031	4470	0.557
GST-dsRBD3	35.118	48820	1.390
GST-Loqs-0	41.204	51800	1.257
GST-Loqs-1	40.020	48820	1.220
GST-Loqs-2	38.952	48820	1.253
GST-Loqs-3	37.795	48820	1.292
GST-Loqs-4	36.844	48820	1.325
GST-Loqs-linker	33.305	47330	1.421
GFP-dmDcr-1	276.291	238140	0.862
dmDcr-1	256.564	226220	0.882
TRBP-3	38.302	52830	1.379
TRBP-2	40.572	52830	1.302
TRBP-1	41.655	52830	1.268
Brat-NHL	31.797	16390	0.515

4.2 Preparation of eukaryotic cellular lysates

For the lysis of mammalian cells, three different protocols were used.

Detergent based lysis

For the detergent based lysis, cells from a 15 cm culture plate were washed with PBS and pelleted (500 x g, 5 min, 4 °C). The pellet was resuspended in 1 mL of a detergent based lysis buffer (see Material Chapter 1.8) supplemented with 1 mM AEBSF, 1 mM DTT and 1 mM NaF and incubated for 20 min on ice. The lysate was cleared by centrifugation (17,000 x g, 20 min, 4 °C) and the supernatant was collected for further applications.

Lysis by hypertonic swelling

To generate the lysate for the FCCS loading assays, a modified protocol from Dignam et al., 1983b and Martinez et al., 2002 was used. Therefore cells were harvested and pelleted (500 x g, 5 min and 4 °C) after washing with PBS. After washing, the cells were gently resuspended in Roeder A buffer (5 mL per 1 mL of packed cell volume) and incubated for 10 min on ice. Cells were pelleted for a second time and resuspended in 2mL Roeder A buffer per 1 mL cell pellet and transferred to a glass dounce homogenizer (B type pestle) and homogenized with ten strokes. The lysate was then pre-cleared by centrifugation (2000 x g, 10 min, 4°C) and the supernatant was applied to a 5 mL polycarbonate tube and mounted to a TST 55.5 rotor (Kontron Instrumnets) and centrifuged for 60 min at 100,000 x g at 4°C (Kontron Centrikon T-1170). Prior freezing in liquid nitrogen, 2 mM $MgCl_2$, 100 mM KCl and 5 % glycerol were added.

Lysis by sonication

For the lysis via sonication cell were harvested from the culture plate and washed with PBS. After pelleting of the cells (500 x g, 5 min, 4 °C) cells were resuspended in an appropriate volume of PBS supplemented with 1 mM AEBSF, 1 mM DTT and 1 mM NaF. Cells were lysed via sonication (Pulse 50, Intensity 7, 30 strokes) lysate was clarified via centrifugation (20800 x g, 15 min, 4 °C).

4.3 GST-pulldown assay

For each pulldown sample, 40 µL of Glutathione 4 Fast Flow Sepharose (GE Healthcare) were washed with PBS twice. Recombinant GST-bait proteins (300 µg each) were added to the beads and incubated for 1 h at 4 °C. Then the supernatant was removed and beads were washed three times with PBS and 1 mL of SF21 lysate containing the respective Dicer protein. For the dsRBD3-only constructs, a threefold more concentrated lysate was applied to the coupled beads. After incubation for 3 h at 4°C while rotating, the supernatant was removed and the beads were washed once with PBS supplemented with 1 mM DTT and 1 mM AEBSF and then three times with a wash buffer containing 25 mM Tris, pH 7.5, 350 mM NaCl, 1 mM $MgCl_2$, 1 mM DTT, 1 mM AEBSF and 0.1 % NP40. Samples, used for MA-analysis, additionally contained 1 mM NaF as an inhibitor for phosphatases. The elution was performed with 40 µL elution buffer (GST B) for 30 min at RT and the eluate was supplemented with Laemmli sample buffer (5 fold concentrated). Pulldown samples for MS were directly eluted with 2x Laemmli buffer at 95 °C for 10 min. Proteins were separated on an 8 % SDS-poly-acrylamide gel and Western blotting was performed as described below. As a loading control, 1% of each elution sample was separated on a 10% SDS–polyacrylamide gel and Coomassie-stained.

4.4 Dicer rescue assay

For the dicer rescue assay, 800,000 Dicer$^{-/-}$ MEF-cells were seeded to a 6-well plate and grown overnight. Then cells were transfected with Lipofectamine LTX, using 0.5 µg of VP5-Loqs and 2 µg of VP5-dmDcr-1 as described in Methods Chapter 5.1. For the Loqs-L429R mutant, 1 µg of plasmid was transfected due to a lower expression-level. Cells were transfected in a 6 well plate and incubated for 6 h. Then cells were split to a 10 cm cell culture dish and after incubation for 2 d, cells were harvested. Sixty-five percent of the cells were used for Western blot analysis and 35% for RNA extraction. Cells for Western samples were resuspended in 80 µL PBS supplemented 20 µL 5x Laemmli buffer and lysed via sonication. Western blotting was performed as described below. For Northern blotting analysis, 10 µg of extracted total RNA from the rescued Dicer$^{-/-}$ cells was used and 0.5 µg of the Dicer$^{+/+}$ total RNA as a wt control for let-7a.

4.5 Co-immunoprecipitation assay

HEK293-T cells co-transfected with myc-tagged and FLAG/ HA-tagged constructs. Then cells were harvested and lysed for co-immunoprecipitation (co-IP) 48 h after transfection. For each co-IP, transfected cells from one 15-cm plate were lysed in 1 mL IP-lysis buffer. To exclude an indirect RNA-mediated interaction, 100 µg/mL RNase A (Thermo Scientific) into relevant samples. Lysates were cleared by centrifugation after an incubation time of 15 min at 4°C (additionally 10 min at room temperature upon RNase A treatment). As input samples 1.6 % of the cleared lysates were taken. The residual input was applied to 50 µL Protein-A-Sepharose beads (GE Healthcare), pre-coupled with 2.5 µg anti-c-myc antibody (Sigma Aldrich). The Co-IP was performed for 3 h at 4°C. As control, rabbit IgG (Santa Cruz) was used. After incubation, beads were washed five times with IP wash buffer and once with PBS buffer. The elution of the precipitated protein complexes was performed by adding 50 µL SDS-PAGE loading buffer to the beads followed by heating to 95°C for 5 min. For analysis via Western blot, 40% of each IP-eluate was taken for PAGE.

4.6 SDS PAGE and Western Blot

Samples for SDS-PAGE were mixed with Laemmli Buffer and incubated at 95°C for 5 min. Proteins were separated on SDS-polyacrylamide gels with different percentages (8, 10 and 15%) with constant current (50 mA for normal gel, 30 mA for mini-gel) until the running front reached the bottom of the gel.

For western blotting, proteins were transferred to a Hybond ECL (GE Healthcare) nitrocellulose membrane via semidry- (for low MW proteins) or wet-blotting (for high MW proteins) respectively (2 mA/cm^2 of membrane area, 1 min per 1kDa from protein of interest). After the transfer, the membrane was blocked for 1 h using 5 % milk powder in TBS-T while shaking. For

Immuno-detection primary antibody was diluted in TBS-T (5 % milk, 0.02 % NaN_3) and incubated for 3 hours at RT or o/n at 4 °C. After washing three times with TBS-T, the secondary antibody (5 % milk, 0.02 % NaN_3) was applied for 1 h. After washing (3xTBS-T) the membrane was scanned with the Odyssey IR detection system (Licor). Antibodies used for western blot analysis were listed in Table 4.7.

4.7 Sucrose gradients

Sucrose gradients for separation of different proteins and protein complexes by centrifugation were generated using 15 % to 55 % sucrose in gradient buffer in 14x89 mm polyallomer centrifuge tubes (Beckman, Palo Alto, USA) using the Gradient Master 107ip system (Biocomp, New Brunswick, Canada) and cooled to 4 °C. Then, lysate that was clarified from cell debris by centrifugation at 17000 g for 10 minutes at 4 °C was applied on top of the gradient. Lysates were separated by centrifugation at 30000 rpm in a SW41 rotor for 18 h at 4 °C in an Optima L-90K ultracentrifuge (Beckman Coulter GmbH). A "low" acceleration rate and no brakes were used within the centrifuge-settings. 500 µl fractions were taken manually and used for western blotting in order to check for protein distribution.

5. Cell culture of eukaryotic cells

5.1 Cell culture of mammalian cells

HEK 293T and MEF ($Dcr^{+/+}$ and $Dcr^{-/-}$) cells were cultured in Dulbecco's modified Eagle's medium (DMEM, Sigma-Aldrich) supplemented with 10% FBS (Sigma-Aldrich) and 1% penicillin-streptomycin (Sigma-Aldrich) under standard conditions (37°C, 5% CO_2). After transfection, the antibiotics were omitted from the media to reduce extracellular stress for the cells. The following transfection strategies were used in this thesis.

Transfection with calcium phosphate

The calcium phosphate method was used for "easy to transfect" cells lines like HEK293 cells. The transfection was performed when cells were ~30% confluent. Per 15 cm plate, 5-20 µg of a eukaryotic expression vector encoding for the gene of interest were used and mixed with 123 µL of 2 M $CaCl_2$ solution and filled up to 1 mL with ddH_2O. This solution was further applied drop by drop and while shaking to a second tube containing 1 mL of 2 x HEPES buffered saline (HBS) and incubated at RT for 10 minutes. After incubation, the solution was slowly added to the cells and transient expression was performed for 48 h.

Transfection with Lipofectamine LTX

Lipofectamine LTX (Thermo) was used for "hard to transfect" cells lines like MEF cells. All transfections were performed in 6-well culture plates. For a reverse transfection, 0.8 x 10^6 MEF $Dcr^{-/-}$ cells were used and cultivated overnight. For the forward transfection, which worked better for the MEF $Dcr^{-/-}$ cells, 1.6 x 10^6 cells were used and transfected the same day. For the transfections a total amount of 3 µg of a eukaryotic expression vector were applied to 500 µL of Opti-MEM® I Reduced Serum Media (Thermo Scientific) and supplemented with 2.5 µL of PLUS™ Reagent. After gentle mixing, 8 µL of Lipofectamine were added to the reaction tube and briefly vortexed. The reaction mix was incubated for 15 min at RT and added drop by drop to the cells which were supplied with 2 mL of antibiotic free medium. The transfection was performed for 6-8 h and cells were split to 10 cm culture dishes afterwards.

5.2 Cell culture of insect cells

Drosophila Dmel-2 cells were cultured in Express Five® SFM medium (Thermo Scientific) supplemented with GlutaMAX™ (Thermo Scientific) and Sf21 cells were cultured in Sf-900™ II SFM medium (Thermo Scientific). All insect cells were grown in a tempered cell culture room at 27°C while shaking.

Bac-to-Bac® Baculovirus Expression System (Invitrogen)

Expression of proteins in Sf21 cells with a recombinant baculovirus was performed according to supplier's protocols.

6. Protein-purification

All purification steps of recombinant proteins were performed at 4 °C. Protein concentration was determined spectrophotometrically at 280 nm. GST-tagged proteins were expressed from pGEX-4T-1, His_6-Thioredoxin tagged proteins from pET32a and His tagged proteins from pColdI.

6.1 Expression- and solubility-screens

For the expression of a recombinant protein in *E. coli*, the protein coding sequence (CDS) was inserted into the expression vectors pET32a, pColdI and pGEX4T-1. After transformation, clones were isolated from an agarose plate and cultivated overnight at 37°C. From that culture 300 mL of LB-medium supplemented with selection antibiotics were inoculated and grown to an OD_{600} of 0.6. After induction of protein expression with 1mM IPTG, the culture was split to three 100 mL cultures which were grown at 18, 25 and 37°C respectively. After 2, 4 and 6h (37°C) or 4, 6 and 16 h (18 and 25°C) 1 mL sample was taken and centrifuged for 10 min (5000x g, 4°C). Then the

LB-Medium was removed and the bacterial pellet was re-suspended in 300 µL HisA (His-tagged proteins) or GSTA buffer (GST-tagged proteins) and lysed via sonication (Pulse 50, intensity 3, 45 pulses, RT) and samples were taken before (total) and after (soluble) centrifugation (17000x g, 10 min, 4°C) and supplemented with 1 x Laemmli sample buffer. Expression and solubility were validated via SDS-Gel electrophoresis where 20 µL from each sample were loaded and coomassie stained after running.

6.2 Preparation of bacterial lysates

For the generation of lysates for purification, the best conditions from the initial solubility and expression screens were used. Therefore one colony from a transformed expression-strain containing the respective expression vector was selected and grown over night in 5 mL LB-medium at 37°C. Then an expression culture containing LB-medium with all selection antibiotics was inoculated 1:2000 and grown to an OD_{600} of 0.6 at 37°C. After the induction with 1mM IPTG, expression was performed under the optimal expression-time and –temperature. Bacteria were harvested by centrifugation (5000x g, 10 min, 4°C) and resuspended in 20 mL lysis buffer, containing 1 mM AEBSF and 10 U/mL benzonase (Novagen), per 1 L LB-culture. To prevent oxidation of thiol groups within the recombinant protein, 1 mM DTT was added for all GST-tagged proteins or 5 mM of β-mercaptoethanol for His-tagged proteins. Cells were lysed via sonication (Pulse 50, Intensity 7, 3x3 min on ice) and cell debris were removed with a centrifugation step (40000xg, 40min, 4°C) and the supernatant was filtrated (Pall Life Sciences, Acrodisc 25mm, 0.45 µm membrane) prior loading.

6.3 Purification of His-tagged Loqs-constructs

Crystallization-construct pET32a-Gly_2-dsRBD3

For the generation of high amounts of the recombinant protein, 4 L of an overnight LB-culture were used and after lysate generation, the supernatant was loaded onto the IMAC II column (Material Chapter 1.9) charged with Ni^{2+}. After washing with buffer His-A, elution was conducted with His-B buffer and His_6-trx fusion protein-containing fractions were pooled. The tag was cleaved off during dialysis in presence of a His-tagged TEV protease (1 µg TEV per 100 µg of protein, dialysis overnight at 4°C against His-A buffer). The tag and the protease were removed by a second IMAC II column (Material Chapter 1.9) and the flowthrough was concentrated by ammonium sulfate precipitation (50% w/v), resuspended in SEC 1 buffer and loaded onto a S75 26/600 column (Material Chapter 1.9). Methylation of the protein required for crystallization was performed according to (Walter et al., 2006) and excess of methylation agent was removed with a second S75 26/600 column equilibrated with buffer SEC 2. After desalting to SEC 3 buffer with

HiPrep Desalt 26/60, the final concentration of the methylated protein was adjusted to 15 mg/mL and used for protein crystallization (Results Chapter 2.3).

6.4 Purification of GST-tagged Loqs-constructs

GST-Tagged proteins were loaded onto a 5 mL GSTrap column (GE), which was equilibrated with GST-A buffer prior usage. Unbound proteins were removed using 5 column volumes (cv) GST-A buffer supplemented with 1M NaCl. After 1 cv of GST-A, bound protein was eluted with GST-B buffer. GST fusion protein containing fractions were pooled and applied to a Hiprep 26/10 desalting column and buffer was exchanged to GST-A.

Proteins, which were used for the GST pulldown assay, were supplemented with 5 % glycerol and 1mM DTT, frozen in liquid nitrogen and stored at -80 °C.

For other applications, where the tag was not needed, GST-proteins were digested overnight at 4°C, using a GST-tagged TEV protease. The protease and the GST-tag were removed by a second 5 mL GSTrap column. The flowthrough was pooled and concentrated. Purified proteins were supplemented with 5 % glycerol and 1mM DTT, frozen in liquid nitrogen and stored at -80 °C.

6.5 Purification of Brat-constructs

Brat-constructs were fused to a His_6-ubiquitin-tag using the pHUE vector system as described previously (Baker et al., 2005; Catanzariti et al., 2004). The *Drosophila* Brat-NHL domain was expressed in *Escherichia coli* (*E.coli*) BL21(DE3). A lysate containing the Brat-NHL domain was loaded on the IMAC I (Material Chapter 1.9) column and eluted by buffer HisB. His_6-ubiquitin fusion protein-containing fractions were pooled, and the His_6-ubiquitin moiety was cleaved off by incubation with the Usp2cc enzyme overnight at 4°C in buffer HisB containing 1 mM DTT. The protein solution was subsequently applied to a S75 26/60 column (Material Chapter 1.9) equilibrated with SEB2. Fractions containing highly pure protein were pooled, supplemented with 5 % glycerol and 1mM DTT, frozen in liquid nitrogen and stored at -80 °C.

7. X-ray crystallography

7.1 Crystallization

All crystallization trials were performed using the vapor diffusion method. Thereby a droplet with the purified protein is mixed with reservoir solution containing the precipitant and equilibrated against a higher volume of the reservoir in a sealed reaction chamber. Upon equilibration of the protein solution and the reservoir solution, the protein concentration increases and leads to a destabilization of the solution that ideally results in crystal formation.

Initial crystals were screened using the "sitting drop" method. For this purpose, commercially available crystal screens (Table 4.11) were used, that were generated on statistical analysis of conditions which promoted crystal growth. 70 µl of each reservoir solution from a crystal screen were submitted to a 96 well plate (Hampton Research) and 1 µl of the respective reservoir solution was mixed with 1 µl of the protein solution on a pedestal on top of the reservoir solution. Then the 96 well plate was sealed with Crystal Clear Tape and stored at 4, 20 and 37°C respectively.

Table 4.11: List of different crystal screens.

Crystal screen	Distributor
JBScreen classic 1-10	Jena Bioscience (Jena, Germany)
JBScreen basic 1-4	Jena Bioscience (Jena, Germany)
JBScreen Pentaerythritol	Jena Bioscience (Jena, Germany)
JBScreen Wizard	Jena Bioscience (Jena, Germany)
Crystal screen 1+2	Hampton Research (Aliso Viejo, USA)
Crystal Screen Cryo 1+2	Hampton Research (Aliso Viejo, USA)
Natrix 1+2	Hampton Research (Aliso Viejo, USA)

Cryo protection and preparation of protein crystals

For the storage and measurements of protein crystals, a suitable cryoprotectant needs to be identified. A cryoprotectant, when added to the crystal and crystallization reagent, is able to generate a glass like surrounding around the crystal upon cooling to cryogenic temperature. This surrounding prevents the formation of ice rings and cryo-damage to the crystal. As cryoprotectants, glycerol, Methylpentanediol (MPD), low-molecular-weight polyethylene glycols (PEG), alcohols and sugars can be used. For all crystallographic experiments from this thesis, glycerol worked best as cryoprotectant.

After the identification of a suitable cryoprotectant, the crystals were transferred from the respective drop using cryo-loops (Hampton Research) and briefly soaked in cryo-solution, containing the mother liquor with the cryoprotectant. After incubation in cryo-solution, crystals were loaded into a cryo-loop and flash frozen in liquid nitrogen. Until the measurement at the synchrotron, crystals were stored in a Dewar flask, containing liquid nitrogen.

7.2 Incorporation of SeMet into proteins

For the preparation of SeMet-derivatized proteins, *E.coli* BL21(DE3) cells were transformed with a respective expression vector and grown on LB plates, containing the respective antibiotic(s), at 37°C overnight. A single colony from the plate was used for the inoculation of 100 ml LB medium (with antibiotics) and grown overnight at 37°C. After incubation, cells were harvested

(3000xg, 10 min and 4°C) and resuspended in 100 ml M9 medium (with antibiotics) (Material chapter 1.7), which was used for the inoculation (~1:50 v/v) of 2 L M9 medium (with antibiotics). The culture was grown to an OD_{600} of 0.6 and supplemented with 125 mg/l L-lysine (in H_2O), 100 mg/l L-phenylalanine (in 0.5 M HCl), 100 mg/l L-threonine (in 0.1 M HCl), 50 mg/l L-isoleucine (in 0.1 M HCl), 50 mg/l L-leucine (in 0.5 M HCl), 50 mg/l L-valine (in 0.5 M HCl) and 60 mg/l L-SeMet (in 0.1 M HCl) (modified from Van Duyne et al., 1993). The added HCl was neutralized by the addition of an equal amount on 2 M NaOH solution and cells were incubated for another 15 min at 37°C. Protein expression was induce with 1 mM IPTG and cells were grown overnight at 18°C. To prevent oxidation of the SeMet side chain, reducing agents, 10 mM β-mercaptoethanol for His_6-containing proteins or 1 mM DTT for all other proteins, were added to every chromatographic step upon purification.

7.3 Data-collection and -processing

All data from the crystals were collected at BESSY (beamline BL14.2) in Berlin. Therefore, crystals were rotated in 0.5 ° steps and a picture after each rotation was taken. The length of exposure for each image was ~ 5 sec. For the measurement of native protein crystals, a wavelength of ca. 1 Å (ca. 12.4 keV) was used. For the measurements of MAD datasets, a fluorescence scan was performed to obtain suitable wavelengths for the peak-, inflection- and high remote-datasets (See Chapter 6.4.2).

All data from this thesis were processed using the program XDS (Kabsch, 2010) together with the XDS-integrated programs XSCALE and XDSCONV, which leads to scaling and reduction of the data. The processing results in the cell parameters of the asymmetric unit cell (ASU). Additionally the Wilson-B-factor (temperature factor) is obtained using the program XSCALE.

After the identification of the cell parameters using XDS, the Matthews coefficient can be calculated using MATTHEWS_COEFF (ccp4-suite (Collaborative Computational Project, 1994)), resulting in the solvent content and the number of protomers within the ASU.

7.4 Phasing

7.4.1 Molecular replacement

Molecular replacement (MR) can be used for the determination of the phase problem, when the structure of the crystallized protein has been solved or if a structure of a homologous protein exists. Thereby the known structure-model of the crystallized protein can be used as a search model. Through three dimensional rotation- and translation-operations, the search model is positioned within the electron density of the ASU, obtained for the newly crystallized protein. For this thesis, the program PHASER (McCoy, 2007) was used, which is based on the maximum likelihood method. Thereby, Z-scores above 6.0 can be regarded as reliable solutions.

7.4.2 Experimental phasing

In cases where MR does not provide a solution, the phase problem has to be solved experimentally. This can be achieved by the incorporation of an electron rich heavy metal atom either by soaking protein crystals in a solution containing the heavy metal atom, or, as it is the case for this thesis, by incorporation of a heavy metal atom-derivatized protein, in this case selenium (SeMet). The anomalous diffraction from the heavy metal atoms, obtained by three different wavelengths (peak-, inflection- and high remote) is needed for the determination of the atomic positions of the heavy metal atoms within the ASU. For the calculation the and refinement of the heavy metal atom-positions, the programs SHARP and autoSHARP (Vonrhein et al., 2007) were used, respectively.

7.5 Model building

All model building was performed manually using the program COOT (Emsley et al., 2010). For refinement of the model and to reduce model bias, the program PHENIX.*refine* (Afonine et al., 2012) was used. Thereby, non-crystallographic symmetry (NCS) restraints and simulated annealing were used. At the final stages of the Loqs dsRBD3, TLS-refinement was carried out. The model of the bound consensus front RNA at the Brat NHL-domain was improved by rebuilding using ERRASER (Chou et al., 2012).

8. Fluorescence spectroscopic methods

8.1 Fluorescence cross-correlation spectroscopy FCCS

All FCCS measurements were carried out in the laboratory of Dr. Stefan Hannus and Dr. Frank Becker (Intana Bioscience, Munich). As an instrumental setup, a ConfoCor2 FCS unit connected to an Axiovert 100M stand equipped with a C-Apochromat NA 1.2 40-fold water immersion lens (Carl Zeiss, Jena, Germany). Excitation of fluorescing compounds was performed with a 488nm laser line of an argon-ion laser and with a 633 helium-neon laser respectively. Emitted photons were directed over an HFT 488/633 and NFT 635/IR beam splitter and detected after passing a BP 505-550 filter or a LP 650 filter by an avalanche photo diode. Samples were distributed to a 384-well glass bottom assay plate (SWISSCI AG, Neuheim, Switzerland) and acquisition was performed 5-8 times for 8-10 sec per data point.

For titration experiments, the GFP labeled compound was diluted either with empty S100 lysate containing no GFP labeled proteins or in case of recombinant proteins with PBS containing 0.2 % BSA to prevent unspecific binding to the surfaces of the 384-well plate. The protein solution was dispensed as 20 µl aliquots and an Alexa 647 (Loqs-3) or Cy5 (siRNA) labeled component was added with the highest concentration to the first well. From that, serial dilutions from well to well

were carried out to generate a titration curve with decreasing concentrations. After incubation at RT for different time points, the FCCS measurements were done and fluorescence fluctuations were auto- and cross-correlated. A fitting-algorithm was employed accounting for two diffusing species for one fluorescently labeled compound and one diffusing species for the other one which was determined before by measuring the diffusion constant of the single compound in solution. The K_d value was calculated on the basis of free and bound fractions of both labelled interactors and the concentration of doubly labelled complexes.

For the competition assays, the K_d value of one labeled reference component was determined as described above and a dual labeled complex was formed and distributed as a 20 µL aliquot to several wells of the 384 well plate where the competition assay is performed. A serial dilution for the unlabeled competitor was prepared and added to each well. After incubation at RT for 10 min, the FCCS measurements were performed and IC50 values were calculated by plotting of the concentration from the dual labeled complex against the concentration of the competitor. From the resulting IC50 values, the K_i values could be obtained by application of the Cheng Prusoff equation.

8.2 Single-molecule co-IP experiments

Single-molecule co-IP assays were carried out in a custom-built flow chambers based on fused silica slides, passivated with polyethylene glycol (PEG). Flow chambers were prepared and assembled as described previously (Gietl et al., 2014). For fluorescence measurements on single immobilized proteins, the flow chamber was filled with 0.1 mg/mL NeutrAvidin (Thermo Scientific) and incubated for 5 min and washed with PBS. Afterwards, the chamber was incubated with 15 nM biotinylated anti-GFP-antibody (Abcam) in PBS and washed again with PBS supplemented with 0.5 mg/mL BSA (Roche). In order to test for unspecific binding of Loqs-3AF647, the chamber was incubated with Loqs-3^{AF647} in increasing concentrations (50 to 200 pM) for 5 min, washed with PBS/BSA and incubated for 5 min with oxygen scavenging buffer (Rasnik et al., 2006). In order to measure the oligomerization state of Loqs-3^{AF647} in complex with GFP-dmDcr1, the chamber was incubated with preformed 50-200 pM GFP-dmDicer-1/Loqs-3^{AF647} complexes in PBS/BSA for 5 min, washed with PBS/BSA, and incubated for additional 5 min with oxygen scavenging buffer. Prior incubation of the chamber, the GFP-dmDcr1/Loqs-3^{AF647} complex was pre-formed by incubating GFP-dmDcr1 (final concentration 1.1 µM) with Loqs-^{3AF647} (final concentration 1.6 µM) for 2 h on ice before dilution with PBS/BSA. This procedure yielded a completely saturated GFP-dmDcr1 surface due to low-affinity interaction between Dicer-Loqs and immobilized Loqs-3^{AF647} density of approximately one molecule per 4 μm^2. Afterwards, the flow chamber was flushed and incubated for 5 min with 1× PBS/BSA containing the glucose oxidase/catalase oxygen scavenging system. As a control, free AF647 dye (200pM)

was flushed over an antibody-GFP-dmDcr-1 surface and incubated for 5 min followed by oxygen removal. Single-molecule fluorescence measurements were performed on a homebuilt prism-type total internal reflection (TIRF) setup based on a Leica DMi8 inverse research microscope. Fluorophores were exited with a 488 nm solid-state laser (CoherentOBIS) with a power of 10 mW and 637 nm diode laser (Coherent OBIS, clean-up filter ZET 635/10, AHF Göttingen) with a power of 50 mW. The fluorescence was collected by a Leica HC PL Apo 63x N.A. 1.20 water immersion objective and split by wavelength with a dichroic mirror (HC BS 560, AHF Göttingen) into two detection channels that were further filtered with a 635-nm long-pass filter (LP Edge Basic, AHF Göttingen) in the red detection channel. Both detection channels were recorded by one EMCCD camera (Andor IX on Ultra 897, EM-gain 20, frame rate 10 Hz, 200–300 frames) in a dual-view configuration (TripleSplit, Cairn Research). The videos were analyzed employing the iSMS software (Preus et al., 2015). Molecule spots were detected using a threshold of 200 for Alexa647 spots. Single molecule data were acquired as the average number of Loqs-3^{AF647} fluorescent molecules per imaging area (6800 μm2) as shown in the histograms. The error bars represent standard deviation of the mean values from 11 imaging areas. The number of fluorescence photo bleaching steps of Alexa 647-labeled Loquacious was determined for detected spots with a fluorescence intensity of at least 600 (arbitrary units) for single-step bleaching and 1200 for double-step bleaching. The number of molecules showing one or two bleaching steps was accumulated from four independent experiments to obtain the stoichiometry of the complex.

Bibliography

Afonine, P. V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., Adams, P.D., 2012. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. Sect. D Biol. Crystallogr. 68, 352–367. doi:10.1107/S0907444912001308

Aksentijevich, I., Centola, M., Deng, Z., Sood, R., Balow J.E., J., Wood, G., Zaks, N., Mansfield, E., Chen, X., Eisenberg, S., Vedula, A., Shafran, N., Raben, N., Pras, E., Pras, M., Kastner, D.L., Blake, T., Baxevanis, A.D., Robbins, C., Krizman, D., Collins, F.S., Liu, P.P., Chen, X., Shohat, M., Hamon, M., Kahan, T., Cercek, A., Rotter, J.I., Fischel-Ghodsian, N., Richards, N., Shelton, D.A., Gumucio, D., Yokoyama, Y., Mangelsdorf, M., Orsborn, A., Richards, R.I., Ricke, D.O., Buckingham, J.M., Moyzis, R.K., Deaven, L.L., Doggett, N.A., 1997. Ancient missense mutations in a new member of the RoRet gene family are likely to cause familial Mediterranean fever. Cell 90, 797–807. doi:10.1016/S0092-8674(00)80539-5

Albertini, A. a V, Wernimont, A.K., Muziol, T., Ravelli, R.B.G., Clapier, C.R., Schoehn, G., Weissenhorn, W., Ruigrok, R.W.H., 2006. Crystal structure of the rabies virus nucleoprotein-RNA complex. Science 313, 360–3. doi:10.1126/science.1125280

Allers, J., Shamoo, Y., 2001. Structure-based analysis of protein-RNA interactions using the program ENTANGLE. J. Mol. Biol. 311, 75–86. doi:10.1006/jmbi.2001.4857

Anantharaman, V., Koonin, E. V, Aravind, L., 2002. Comparative genomics and evolution of proteins involved in RNA metabolism. Nucleic Acids Res. 30, 1427–1464. doi:10.1093/nar/30.7.1427

Arama, E., Dickman, D., Kimchie, Z., Shearn, A., Lev, Z., 2000. Mutations in the beta-propeller domain of the Drosophila brain tumor (brat) protein induce neoplasm in the larval brain. Oncogene 19, 3706–16. doi:10.1038/sj.onc.1203706

Auweter, S.D., Fasan, R., Reymond, L., Underwood, J.G., Black, D.L., Pitsch, S., Allain, F.H.-T., 2006. Molecular basis of RNA recognition by the human alternative splicing factor Fox-1. EMBO J. 25, 163–73. doi:10.1038/sj.emboj.7600918

Auweter, S.D., Oberstrass, F.C., Allain, F.H.T., 2006. Sequence-specific binding of single-stranded RNA: Is there a code for recognition? Nucleic Acids Res. 34, 4943–4959. doi:10.1093/nar/gkl620

Auyeung, V.C., Ulitsky, I., McGeary, S.E., Bartel, D.P., 2013. Beyond secondary structure: primary-sequence determinants license pri-miRNA hairpins for processing. Cell 152, 844–58. doi:10.1016/j.cell.2013.01.031

Bacia, K., Schwille, P., 2007. Practical guidelines for dual-color fluorescence cross-correlation spectroscopy. Nat. Protoc. 2, 2842–2856. doi:10.1038/nprot.2007.410

Bailey, T.L., Elkan, C., 1994. Fitting a mixture model by expectation maximization to discover

motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol. 2, 28–36.

Baillat, D., Shiekhattar, R., 2009. Functional dissection of the human TNRC6 (GW182-related) family of proteins. Mol. Cell. Biol. 29, 4144–55. doi:10.1128/MCB.00380-09

Baker, R.T., Catanzariti, A.-M., Karunasekara, Y., Soboleva, T.A., Sharwood, R., Whitney, S., Board, P.G., 2005. Using deubiquitylating enzymes as research tools. Methods Enzymol. 398, 540–54. doi:10.1016/S0076-6879(05)98044-0

Baltz, A.G., Munschauer, M., Schwanhäusser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M., Wyler, E., Bonneau, R., Selbach, M., Dieterich, C., Landthaler, M., 2012. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol. Cell 46, 674–90. doi:10.1016/j.molcel.2012.05.021

Barlow, P.N., Luisi, B., Milner, A., Elliott, M., Everett, R., 1994. Structure of the C3HC4 domain by 1H-nuclear magnetic resonance spectroscopy. A new structural class of zinc-finger. J. Mol. Biol. doi:10.1006/jmbi.1994.1222

Berghammer, H., Auer, B., 1993. "Easypreps": fast and easy plasmid minipreparation for analysis of recombinant clones in E. coli. Biotechniques 14, 524, 528.

Bernstein, E., Caudy, a a, Hammond, S.M., Hannon, G.J., 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409, 363–366. doi:10.1038/35053110

Betschinger, J., Mechtler, K., Knoblich, J.A., 2006. Asymmetric Segregation of the Tumor Suppressor Brat Regulates Self-Renewal in Drosophila Neural Stem Cells. Cell 124, 1241–1253. doi:10.1016/j.cell.2006.01.038

Billy, E., Brondani, V., Zhang, H., Müller, U., Filipowicz, W., 2001. Specific interference with gene expression induced by long, double-stranded RNA in mouse embryonal teratocarcinoma cell lines. Proc. Natl. Acad. Sci. U. S. A. 98, 14428–33. doi:10.1073/pnas.261562698

Bogerd, H.P., Whisnant, A.W., Kennedy, E.M., Flores, O., Cullen, B.R., 2014. Derivation and characterization of Dicer- and microRNA-deficient human cells. RNA (New York, NY) 20, 923–937. doi:10.1261/rna.044545.114.1

Bohnsack, M.T., Czaplinski, K., Gorlich, D., 2004. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. RNA 10, 185–91. doi:10.1261/rna.5167604

Borden, K.L., Lally, J.M., Martin, S.R., O'Reilly, N.J., Etkin, L.D., Freemont, P.S., 1995. Novel topology of a zinc-binding domain from a protein involved in regulating early Xenopus development. EMBO J. 14, 5947–5956.

Bork, P., Holm, L., Sander, C., 1994. The Immunoglobulin Fold. J. Mol. Biol. 242, 309–320. doi:10.1006/jmbi.1994.1582

Braddock, D.T., Louis, J.M., Baber, J.L., Levens, D., Clore, G.M., 2002. Structure and dynamics

of KH domains from FBP bound to single-stranded DNA. Nature 415, 1051–6. doi:10.1038/4151051a

Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248–54.

Brand, A.H., Livesey, F.J., 2011. Neural Stem Cell Biology in Vertebrates and Invertebrates: More Alike than Different? Neuron 70, 719–729. doi:10.1016/j.neuron.2011.05.016

Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., Hannon, G.J., 2007. Discrete Small RNA-Generating Loci as Master Regulators of Transposon Activity in Drosophila. Cell 128, 1089–1103. doi:10.1016/j.cell.2007.01.043

Budhidarmo, R., Nakatani, Y., Day, C.L., 2012. RINGs hold the key to ubiquitin transfer. Trends Biochem. Sci. 37, 58–65. doi:10.1016/j.tibs.2011.11.001

Bycroft, M., Grünert, S., Murzin, A.G., Proctor, M., St Johnston, D., 1995. NMR solution structure of a dsRNA binding domain from Drosophila staufen protein reveals homology to the N-terminal domain of ribosomal protein S5. EMBO J. 14, 3563–71.

Calero, G., Wilson, K.F., Ly, T., Rios-Steiner, J.L., Clardy, J.C., Cerione, R.A., 2002. Structural basis of m7GpppG binding to the nuclear cap-binding protein complex. Nat. Struct. Biol. 9, 912–7. doi:10.1038/nsb874

Caplen, N.J., Fleenor, J., Fire, a, Morgan, R. a, 2000. dsRNA-mediated gene silencing in cultured Drosophila cells: a tissue culture model for the analysis of RNA interference. Gene 252, 95–105. doi:S0378-1119(00)00224-9 [pii]

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M., Krijgsveld, J., Hentze, M.W., 2012. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell 149, 1393–406. doi:10.1016/j.cell.2012.04.031

Catanzariti, A.-M., Soboleva, T. a, Jans, D. a, Board, P.G., Baker, R.T., 2004. An efficient system for high-level expression and easy purification of authentic recombinant proteins. Protein Sci. 13, 1331–9. doi:10.1110/ps.04618904

Chang, H., Lim, J., Ha, M., Kim, V.N., 2014. TAIL-seq: genome-wide determination of poly(A) tail length and 3’ end modifications. Mol. Cell 53, 1044–52. doi:10.1016/j.molcel.2014.02.007

Chekulaeva, M., Filipowicz, W., Parker, R., 2009. Multiple independent domains of dGW182 function in miRNA-mediated repression in Drosophila. RNA 15, 794–803. doi:10.1261/rna.1364909

Chekulaeva, M., Mathys, H., Zipprich, J.T., Attig, J., Colic, M., Parker, R., Filipowicz, W., 2011. miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs. Nat. Struct. Mol. Biol. 18, 1218–26. doi:10.1038/nsmb.2166

Chen, M., Manley, J.L., 2009. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. Nat. Rev. Mol. Cell Biol. 10, 741–54. doi:10.1038/nrm2777

Chen, Y., Boland, A., Kuzuoğlu-Öztürk, D., Bawankar, P., Loh, B., Chang, C.-T., Weichenrieder, O., Izaurralde, E., 2014. A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing. Mol. Cell 54, 737–50. doi:10.1016/j.molcel.2014.03.034

Chen, Y., Kortemme, T., Robertson, T., Baker, D., Varani, G., 2004. A new hydrogen-bonding potential for the design of protein-RNA interactions predicts specific contacts and discriminates decoys. Nucleic Acids Res. 32, 5147–5162. doi:10.1093/nar/gkh785

Chendrimada, T.P., Gregory, R.I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., Shiekhattar, R., 2005. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature 436, 740–4. doi:10.1038/nature03868

Cho, P.F., Gamberi, C., Cho-Park, Y., Cho-Park, I.B., Lasko, P., Sonenberg, N., 2006. Cap-Dependent Translational Inhibition Establishes Two Opposing Morphogen Gradients in Drosophila Embryos. Curr. Biol. 16, 2035–2041. doi:10.1016/j.cub.2006.08.093

Chong, M.M.W., Zhang, G., Cheloufi, S., Neubert, T.A., Hannon, G.J., Littman, D.R., 2010. Canonical and alternate functions of the microRNA biogenesis machinery. Genes Dev. 24, 1951–60. doi:10.1101/gad.1953310

Chou, F.-C., Sripakdeevong, P., Dibrov, S.M., Hermann, T., Das, R., 2012. Correcting pervasive errors in RNA crystallography through enumerative structure prediction. Nat. Methods 10, 74–76. doi:10.1038/nmeth.2262

Christie, M., Boland, A., Huntzinger, E., Weichenrieder, O., Izaurralde, E., 2013. Structure of the PAN3 pseudokinase reveals the basis for interactions with the PAN2 deadenylase and the GW182 proteins. Mol. Cell 51, 360–373. doi:10.1016/j.molcel.2013.07.011

Collaborative Computational Project, N. 4, 1994. The CCP4 suite: programs for protein crystallography. Acta Crystallogr. Sect. D Biol. Crystallogr. 50, 760–763. doi:10.1107/S0907444994003112

Cox, D.N., Chao, A., Lin, H., 2000. piwi encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. Development 127, 503–14. doi:10631171

Cozzi, F., Siegel, J.S., 1995. Interaction between stacked aryl groups in 1,8-diarylnaphthalenes: Dominance of polar/π over charge-transfer effects. Pure Appl. Chem. 67, 683–689. doi:10.1351/pac199567050683

Crichlow, G. V, Zhou, H., Hsiao, H.-H., Frederick, K.B., Debrosse, M., Yang, Y., Folta-Stogniew, E.J., Chung, H.-J., Fan, C., De la Cruz, E.M., Levens, D., Lolis, E., Braddock, D., 2008. Dimerization of FIR upon FUSE DNA binding suggests a mechanism of c-myc

inhibition. EMBO J. 27, 277–89. doi:10.1038/sj.emboj.7601936

Czech, B., Malone, C.D., Zhou, R., Stark, A., Schlingeheyde, C., Dus, M., Perrimon, N., Kellis, M., Wohlschlegel, J.A., Sachidanandam, R., Hannon, G.J., Brennecke, J., 2008. An endogenous small interfering RNA pathway in Drosophila. Nature 453, 798–802. doi:10.1038/nature07007

Das, P.P., Bagijn, M.P., Goldstein, L.D., Woolford, J.R., Lehrbach, N.J., Sapetschnig, A., Buhecha, H.R., Gilchrist, M.J., Howe, K.L., Stark, R., Matthews, N., Berezikov, E., Ketting, R.F., Tavaré, S., Miska, E.A., 2008. Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the Caenorhabditis elegans germline. Mol. Cell 31, 79–90. doi:10.1016/j.molcel.2008.06.003

Dautry, F., Ribet, C., Buccheri, M.-A., 2007. RNA interference in mammalian cells. J. Soc. Biol. 201, 339–348. doi:10.1126/science.1241930

Deardorff, J. a, Sachs, a B., 1997. Differential effects of aromatic and charged residue substitutions in the RNA binding domains of the yeast poly(A)-binding protein. J. Mol. Biol. 269, 67–81. doi:10.1006/jmbi.1997.1013

Deechongkit, S., Nguyen, H., Powers, E.T., Dawson, P.E., Gruebele, M., Kelly, J.W., 2004. Context-dependent contributions of backbone hydrogen bonding to beta-sheet folding energetics. Nature 430, 101–105. doi:10.1038/nature02611

Deerberg, A., Willkomm, S., Restle, T., 2013. Minimal mechanistic model of siRNA-dependent target RNA slicing by recombinant human Argonaute 2 protein. Proc. Natl. Acad. Sci. U. S. A. 110, 17850–5. doi:10.1073/pnas.1217838110

Deng, W., Lin, H., 2002. Miwi, a Murine Homolog of Piwi, Encodes a Cytoplasmic Protein Essential for Spermatogenesis. Dev. Cell 2, 819–830. doi:10.1016/S1534-5807(02)00165-X

Denli, A.M., Tops, B.B.J., Plasterk, R.H. a, Ketting, R.F., Hannon, G.J., 2004. Processing of primary microRNAs by the Microprocessor complex. Nature 432, 231–5. doi:10.1038/nature03049

Dieci, G., Fiorino, G., Castelnuovo, M., Teichmann, M., Pagano, A., 2007. The expanding RNA polymerase III transcriptome. Trends Genet. 23, 614–622. doi:10.1016/j.tig.2007.09.001

Dignam, J.D., Lebovitz, R.M., Roeder, R.G., 1983a. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. 11, 1475–1489. doi:10.1093/nar/11.5.1475

Dignam, J.D., Lebovitz, R.M., Roeder, R.G., 1983b. Accurate transcription initiation by RNA polymerase D in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. 11, 1475–1489. doi:10.1093/nar/gkq840

Drozdetskiy, A., Cole, C., Procter, J., Barton, G.J., 2015. JPred4: a protein secondary structure prediction server. Nucleic Acids Res. 43, W389-94. doi:10.1093/nar/gkv332

Dueck, A., Meister, G., 2014. Assembly and function of small RNA - argonaute protein

complexes. Biol. Chem. 395, 611–29. doi:10.1515/hsz-2014-0116

Dueck, A., Ziegler, C., Eichner, A., Berezikov, E., Meister, G., 2012. microRNAs associated with the different human Argonaute proteins. Nucleic Acids Res. 40, 9850–62. doi:10.1093/nar/gks705

Edwards, T.A., Pyle, S.E., Wharton, R.P., Aggarwal, A.K., 2001. Structure of pumilio reveals similarity between RNA and peptide binding motifs. Cell 105, 281–289. doi:10.1016/S0092-8674(01)00318-X

Edwards, T. a, Wilkinson, B.D., Wharton, R.P., Aggarwal, A.K., 2003. Model of the brain tumor-Pumilio translation repressor complex. Genes Dev. 17, 2508–13. doi:10.1101/gad.1119403

Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., Tuschl, T., 2001. Duplexes of 21-nucleotide RNAs meditae RNA interference in cultured mammalian cells. Nature 411, 494–498. doi:10.1038/35078107

Elkayam, E., Kuhn, C.D., Tocilj, A., Haase, A.D., Greene, E.M., Hannon, G.J., Joshua-Tor, L., 2012. The structure of human argonaute-2 in complex with miR-20a. Cell 150, 100–110. doi:10.1016/j.cell.2012.05.017

Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Acta Crystallogr. D. Biol. Crystallogr. 66, 486–501. doi:10.1107/S0907444910007493

Fabian, M.R., Cieplak, M.K., Frank, F., Morita, M., Green, J., Srikumar, T., Nagar, B., Yamamoto, T., Raught, B., Duchaine, T.F., Sonenberg, N., 2012. miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4–NOT. Nat. Struct. Mol. Biol. 19, 1211–1217. doi:10.1038/nsmb0312-364c

Faller, M., Matsunaga, M., Yin, S., Loo, J. a, Guo, F., 2007. Heme is involved in microRNA processing. Nat. Struct. Mol. Biol. 14, 23–9. doi:10.1038/nsmb1182

Ferrandon, D., Elphick, L., Nüsslein-Volhard, C., St Johnston, D., 1994. Staufen protein associates with the 3'UTR of bicoid mRNA to form particles that move in a microtubule-dependent manner. Cell 79, 1221–32. doi:10.1016/0092-8674(94)90013-2

Ferrandon, D., Koch, I., Westhof, E., Nüsslein-Volhard, C., 1997. RNA-RNA interaction is required for the formation of specific bicoid mRNA 3' UTR-STAUFEN ribonucleoprotein particles. EMBO J. 16, 1751–8. doi:10.1093/emboj/16.7.1751

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., Mello, C.C., 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391, 806–811. doi:10.1038/35888

Förstemann, K., Horwich, M.D., Wee, L., Tomari, Y., Zamore, P.D., 2007. Drosophila microRNAs are sorted into functionally distinct argonaute complexes after production by dicer-1. Cell 130, 287–97. doi:10.1016/j.cell.2007.05.056

Förstemann, K., Tomari, Y., Du, T., Vagin, V. V, Denli, A.M., Bratu, D.P., Klattenhoff, C., Theurkauf, W.E., Zamore, P.D., 2005. Normal microRNA maturation and germ-line stem

cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. PLoS Biol. 3, e236. doi:10.1371/journal.pbio.0030236

Frank, D.J., Edgar, B.A., Roth, M.B., 2002. The Drosophila melanogaster gene brain tumor negatively regulates cell growth and ribosomal RNA synthesis. Development 129, 399–407.

Frank, D.J., Roth, M.B., 1998. ncl-1 is required for the regulation of cell size and ribosomal RNA synthesis in Caenorhabditis elegans. J Cell Biol 140, 1321–1329.

Freemont, P.S., 1993. The RING Finger: A Novel Protein Sequence Motif Related to the Zinc Finger. Ann. New York Acad. Sci. 684, 174–192.

Friedman, J.R., Fredericks, W.J., Jensen, D.E., Speicher, D.W., Huang, X.P., Neilson, E.G., Rauscher, F.J., 1996. KAP-1, a novel corepressor for the highly conserved KRAB repression domain. Genes Dev. 10, 2067–2078. doi:DOI 10.1101/gad.10.16.2067

Friedman, R.C., Farh, K.K.H., Burge, C.B., Bartel, D.P., 2009. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res. 19, 92–105. doi:10.1101/gr.082701.108

Frohn, A., Eberl, H.C., Stöhr, J., Glasmacher, E., Rüdel, S., Heissmeyer, V., Mann, M., Meister, G., 2012. Dicer-dependent and -independent Argonaute2 protein interaction networks in mammalian cells. Mol. Cell. Proteomics 11, 1442–56. doi:10.1074/mcp.M112.017756

Fukunaga, R., Han, B.W., Hung, J.H., Xu, J., Weng, Z., Zamore, P.D., 2012. Dicer partner proteins tune the length of mature miRNAs in flies and mammals. Cell 151, 533–546. doi:10.1016/j.cell.2012.09.027

Gallivan, J.P., Dougherty, D.A., 1999. Cation-pi interactions in structural biology. Proc. Natl. Acad. Sci. 96, 9459–9464. doi:10.1073/pnas.96.17.9459

Gao, M., Wei, W., Li, M.-M., Wu, Y.-S., Ba, Z., Jin, K.-X., Li, M.-M., Liao, Y.-Q., Adhikari, S., Chong, Z., Zhang, T., Guo, C.-X., Tang, T.-S., Zhu, B.-T., Xu, X.-Z., Mailand, N., Yang, Y.-G., Qi, Y., Rendtlew Danielsen, J.M., 2014. Ago2 facilitates Rad51 recruitment and DNA double-strand break repair by homologous recombination. Cell Res. 24, 532–41. doi:10.1038/cr.2014.36

Gerber, A.P., Luschnig, S., Krasnow, M. a, Brown, P.O., Herschlag, D., 2006. Genome-wide identification of mRNAs associated with the translational regulator PUMILIO in Drosophila melanogaster. Proc. Natl. Acad. Sci. U. S. A. 103, 4487–92. doi:10.1073/pnas.0509260103

Gerstberger, S., Hafner, M., Tuschl, T., 2014. A census of human RNA-binding proteins. Nat. Rev. Genet. 15, 829–845. doi:10.1038/nrg3813

Ghildiyal, M., Zamore, P.D., 2009. Small silencing RNAs: an expanding universe. Nat. Rev. Genet. 10, 94–108. doi:10.1038/nrg2504

Gietl, A., Holzmeister, P., Blombach, F., Schulz, S., von Voithenberg, L.V., Lamb, D.C., Werner, F., Tinnefeld, P., Grohmann, D., 2014. Eukaryotic and archaeal TBP and TFB/TF(II)B follow different promoter DNA bending pathways. Nucleic Acids Res. 42, 6219–31. doi:10.1093/nar/gku273

Giot, L., Bader, J.S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y.L., Ooi, C.E., Godwin, B., Vitols, E., Vijayadamodar, G., Pochart, P., Machineni, H., Welsh, M., Kong, Y., Zerhusen, B., Malcolm, R., Varrone, Z., Collis, A., Minto, M., Burgess, S., McDaniel, L., Stimpson, E., Spriggs, F., Williams, J., Neurath, K., Ioime, N., Agee, M., Voss, E., Furtak, K., Renzulli, R., Aanensen, N., Carrolla, S., Bickelhaupt, E., Lazovatsky, Y., DaSilva, A., Zhong, J., Stanyon, C.A., Finley, R.L., White, K.P., Braverman, M., Jarvie, T., Gold, S., Leach, M., Knight, J., Shimkets, R.A., McKenna, M.P., Chant, J., Rothberg, J.M., 2003. A protein interaction map of Drosophila melanogaster. Science (80-.). 302, 1727–1736. doi:10.1126/science.1090289

Gleghorn, M.L., Gong, C., Kielkopf, C.L., Maquat, L.E., 2013. Staufen1 dimerizes through a conserved motif and a degenerate dsRNA-binding domain to promote mRNA decay. Nat Struct Mol Biol 20, 515–524. doi:10.1038/nsmb.2528

Gómez-López, S., Lerner, R.G., Petritsch, C., 2014. Asymmetric cell division of stem and progenitor cells during homeostasis and cancer. Cell. Mol. Life Sci. 71, 575–97. doi:10.1007/s00018-013-1386-1

Graindorge, A., Carré, C., Gebauer, F., 2013. Sex-lethal promotes nuclear retention of msl2 mRNA via interactions with the STAR protein HOW. Genes Dev. 27, 1421–33. doi:10.1101/gad.214999.113

Gregory, R.I., Yan, K.-P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., Shiekhattar, R., 2004. The Microprocessor complex mediates the genesis of microRNAs. Nature 432, 235–40. doi:10.1038/nature03120

Griffiths-Jones, S., Saini, H.K., Van Dongen, S., Enright, A.J., 2008. miRBase: Tools for microRNA genomics. Nucleic Acids Res. 36, 154–158. doi:10.1093/nar/gkm952

Grishin, N. V, 2001. KH domain: one motif, two folds. Nucleic Acids Res. 29, 638–43. doi:10.1093/nar/29.3.638

Grütter, C., Briand, C., Capitani, G., Mittl, P.R.E., Papin, S., Tschopp, J., Grütter, M.G., 2006. Structure of the PRYSPRY-domain: Implications for autoinflammatory diseases. FEBS Lett. 580, 99–106. doi:10.1016/j.febslet.2005.11.076

Gupta, Y.K., Nair, D.T., Wharton, R.P., Aggarwal, A.K., 2008. Structures of Human Pumilio with Noncognate RNAs Reveal Molecular Mechanisms for Binding Promiscuity. Structure 16, 549–557. doi:10.1016/j.str.2008.01.006

Haase, A.D., Fenoglio, S., Muerdter, F., Guzzardo, P.M., Czech, B., Pappin, D.J., Chen, C., Gordon, A., Hannon, G.J., 2010. Probing the initiation and effector phases of the somatic piRNA pathway in Drosophila. Genes Dev. 24, 2499–2504. doi:10.1101/gad.1968110

Haase, A.D., Jaskiewicz, L., Zhang, H., Lainé, S., Sack, R., Gatignol, A., Filipowicz, W., 2005. TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. EMBO Rep. 6, 961–7. doi:10.1038/sj.embor.7400509

Hall, I.M., Shankaranarayana, G.D., Noma, K.-I., Ayoub, N., Cohen, A., Grewal, S.I.S., 2002. Establishment and maintenance of a heterochromatin domain. Science 297, 2232–7. doi:10.1126/science.1076466

Hammell, C.M., Lubin, I., Boag, P.R., Blackwell, T.K., Ambros, V., 2009. nhl-2 Modulates microRNA activity in Caenorhabditis elegans. Cell 136, 926–38. doi:10.1016/j.cell.2009.01.053

Hamosh, A., Scott, A.F., Amberger, J.S., Bocchini, C.A., McKusick, V.A., 2005. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Res. 33, 514–517. doi:10.1093/nar/gki033

Han, J., Lee, Y., Yeom, K.H., Kim, Y.K., Jin, H., Kim, V.N., 2004. The Drosha-DGCR8 complex in primary microRNA processing. Genes Dev. 18, 3016–3027. doi:10.1101/gad.1262504

Han, J., Lee, Y., Yeom, K.-H., Nam, J.-W., Heo, I., Rhee, J.-K., Sohn, S.Y., Cho, Y., Zhang, B.-T., Kim, V.N., 2006. Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. Cell 125, 887–901. doi:10.1016/j.cell.2006.03.043

Han, X., Du, H., Massiah, M.A., 2011. Detection and characterization of the in vitro E3 ligase activity of the human MID1 protein. J. Mol. Biol. 407, 505–520. doi:10.1016/j.jmb.2011.01.048

Hartig, J.V., Esslinger, S., Böttcher, R., Saito, K., Förstemann, K., 2009. Endo-siRNAs depend on a new isoform of loquacious and target artificially introduced, high-copy sequences. EMBO J. 28, 2932–2944. doi:10.1038/emboj.2009.220

Hauptmann, J., Dueck, A., Harlander, S., Pfaff, J., Merkl, R., Meister, G., 2013. Turning catalytically inactive human Argonaute proteins into active slicer enzymes. Nat. Struct. Mol. Biol. 20, 814–7. doi:10.1038/nsmb.2577

Hauptmann, J., Schraivogel, D., Bruckmann, A., Manickavel, S., Jakob, L., Eichner, N., Pfaff, J., Urban, M., Sprunck, S., Hafner, M., Tuschl, T., Deutzmann, R., Meister, G., 2015. Biochemical isolation of Argonaute protein complexes by Ago-APP. Proc. Natl. Acad. Sci. U. S. A. 112, 11841–5. doi:10.1073/pnas.1506116112

Herbert, K.M., Sarkar, S.K., Mills, M., Delgado De la Herran, H.C., Neuman, K.C., Steitz, J.A., 2016. A heterotrimer model of the complete Microprocessor complex revealed by single-molecule subunit counting. RNA 22, 175–83. doi:10.1261/rna.054684.115

Hitti, E.G., Sallacz, N.B., Schoft, V.K., Jantsch, M.F., 2004. Oligomerization activity of a double-stranded RNA-binding domain. FEBS Lett. 574, 25–30. doi:10.1016/j.febslet.2004.07.080

Höck, J., Weinmann, L., Ender, C., Rüdel, S., Kremmer, E., Raabe, M., Urlaub, H., Meister, G., 2007. Proteomic and functional analysis of Argonaute-containing mRNA-protein complexes in human cells. EMBO Rep. 8, 1052–60. doi:10.1038/sj.embor.7401088

Holm, L., Rosenstrom, P., 2010. Dali server: conservation mapping in 3D. Nucleic Acids Res. 38, W545–W549. doi:10.1093/nar/gkq366

Houwing, S., Kamminga, L.M., Berezikov, E., Cronembold, D., Girard, A., van den Elst, H., Filippov, D. V., Blaser, H., Raz, E., Moens, C.B., Plasterk, R.H.A., Hannon, G.J., Draper, B.W., Ketting, R.F., 2007. A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. Cell 129, 69–82. doi:10.1016/j.cell.2007.03.026

Huang, Y., Ji, L., Huang, Q., Vassylyev, D.G., Chen, X., Ma, J.-B., 2009. Structural insights into mechanisms of the small RNA methyltransferase HEN1. Nature 461, 823–7. doi:10.1038/nature08433

Hudson, B.P., Martinez-Yamout, M. a, Dyson, H.J., Wright, P.E., 2004. Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d. Nat. Struct. Mol. Biol. 11, 257–264. doi:10.1038/nsmb738

Hülskamp, M., Pfeifle, C., Tautz, D., 1990. A morphogenetic gradient of hunchback protein organizes the expression of the gap genes Krüppel and knirps in the early Drosophila embryo. Nature 346, 577–580. doi:10.1038/346577a0

Iki, T., Yoshikawa, M., Nishikiori, M., Jaudal, M.C., Matsumoto-Yokoyama, E., Mitsuhara, I., Meshi, T., Ishikawa, M., 2010. In vitro assembly of plant RNA-induced silencing complexes facilitated by molecular chaperone HSP90. Mol. Cell 39, 282–291. doi:10.1016/j.molcel.2010.05.014

Im, H.-I., Kenny, P.J., 2012. MicroRNAs in neuronal function and dysfunction. Trends Neurosci. 35, 325–34. doi:10.1016/j.tins.2012.01.004

Inoue, H., Nojima, H., Okayama, H., 1990. High efficiency transformation of Escherichia coli with plasmids. Gene 96, 23–28. doi:0378-1119(90)90336-P [pii]

Ipsaro, J.J., Haase, A.D., Knott, S.R., Joshua-Tor, L., Hannon, G.J., 2012. The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. Nature 491, 279–283. doi:10.1038/nature11502

Ivanov, A. V., Peng, H., Yurchenko, V., Yap, K.L., Negorev, D.G., Schultz, D.C., Psulkowski, E., Fredericks, W.J., White, D.E., Maul, G.G., Sadofsky, M.J., Zhou, M.M., Rauscher, F.J., 2007. PHD Domain-Mediated E3 Ligase Activity Directs Intramolecular Sumoylation of an Adjacent Bromodomain Required for Gene Silencing. Mol. Cell 28, 823–837. doi:10.1016/j.molcel.2007.11.012

Iwasaki, S., Kobayashi, M., Yoda, M., Sakaguchi, Y., Katsuma, S., Suzuki, T., Tomari, Y., 2010. Hsc70/Hsp90 chaperone machinery mediates ATP-dependent RISC loading of small RNA duplexes. Mol. Cell 39, 292–299. doi:10.1016/j.molcel.2010.05.015

Iwasaki, S., Sasaki, H.M., Sakaguchi, Y., Suzuki, T., Tadakuma, H., Tomari, Y., 2015. Defining fundamental steps in the assembly of the Drosophila RNAi enzyme complex. Nature 521, 533–6. doi:10.1038/nature14254

James, L.C., Keeble, A.H., Khan, Z., Rhodes, D.A., Trowsdale, J., 2007. Structural basis for PRYSPRY-mediated tripartite motif (TRIM) protein function. Proc. Natl. Acad. Sci. U. S.

A. 104, 6200–5. doi:10.1073/pnas.0609174104

Jankowsky, E., Harris, M.E., 2015. Specificity and nonspecificity in RNA-protein interactions. Nat Rev Mol Cell Biol 16, 533–544. doi:10.1038/nrm4032

Janowski, B. a, Huffman, K.E., Schwartz, J.C., Ram, R., Nordsell, R., Shames, D.S., Minna, J.D., Corey, D.R., 2006. Involvement of AGO1 and AGO2 in mammalian transcriptional silencing. Nat. Struct. Mol. Biol. 13, 787–92. doi:10.1038/nsmb1140

Janssens, D.H., Lee, C.Y., 2014. It takes two to tango, a dance between the cells of origin and cancer stem cells in the Drosophila larval brain. Semin. Cell Dev. Biol. 28, 63–69. doi:10.1016/j.semcdb.2014.03.006

Jiang, F., Ye, X., Liu, X., Fincher, L., McKearin, D., Liu, Q., 2005. Dicer-1 and R3D1-L catalyze microRNA maturation in Drosophila. Genes Dev. 19, 1674–1679. doi:10.1101/gad.1334005

Joazeiro, C. a, Weissman, a M., 2000. RING finger proteins: mediators of ubiquitin ligase activity. Cell 102, 549–552. doi:10.1016/S0092-8674(00)00077-5

Johnston, M., Geoffroy, M.-C., Sobala, A., Hay, R., Hutvagner, G., 2010. HSP90 protein stabilizes unloaded argonaute complexes and microscopic P-bodies in human cells. Mol. Biol. Cell 21, 1462–9. doi:10.1091/mbc.E09-10-0885

Jonas, S., Izaurralde, E., 2015. Towards a molecular understanding of microRNA-mediated gene silencing. Nat. Rev. Genet. 16, 421–433. doi:10.1038/nrg3965

Kabsch, W., 2010. XDS. Acta Crystallogr. D. Biol. Crystallogr. 66, 125–32. doi:10.1107/S0907444909047337

Kawamata, T., Seitz, H., Tomari, Y., 2009. Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. Nat. Struct. Mol. Biol. 16, 953–60. doi:10.1038/nsmb.1630

Kawaoka, S., Izumi, N., Katsuma, S., Tomari, Y., 2011. 3' End Formation of PIWI-Interacting RNAs In Vitro. Mol. Cell 43, 1015–1022. doi:10.1016/j.molcel.2011.07.029

Keeble, A.H., Khan, Z., Forster, A., James, L.C., 2008. TRIM21 is an IgG receptor that is structurally, thermodynamically, and kinetically conserved. Proc. Natl. Acad. Sci. U. S. A. 105, 6045–50. doi:10.1073/pnas.0800159105

Keren, H., Lev-Maor, G., Ast, G., 2010. Alternative splicing and evolution: diversification, exon definition and function. Nat. Rev. Genet. 11, 345–355. doi:10.1038/nrg2776

Kharrat, a, Macias, M.J., Gibson, T.J., Nilges, M., Pastore, A., 1995. Structure of the dsRNA binding domain of E. coli RNase III. EMBO J. 14, 3572–84.

Khvorova, A., Reynolds, A., Jayasena, S.D., 2003. Functional siRNAs and miRNAs exhibit strand bias. Cell 115, 209–16. doi:10.1016/S0092-8674(03)00801-8

Kim, Y., Yeo, J., Lee, J.H., Cho, J., Seo, D., Kim, J.S., Narry Kim, V., 2014. Deletion of human tarbp2 reveals cellular microRNA targets and cell-cycle function of TRBP. Cell Rep. 9, 1061–1074. doi:10.1016/j.celrep.2014.09.039

Knoblich, J. a, 2010. Asymmetric cell division: recent developments and their implications for tumour biology. Nat. Rev. Mol. Cell Biol. 11, 849–860. doi:10.1038/nrm3010

Knoblich, J. a, Jan, L.Y., Jan, Y.N., 1995. Asymmetric segregation of Numb and Prospero during cell division. Nature 377, 624–7. doi:10.1038/377624a0

Kok, K.H., Ng, M.H.J., Ching, Y.P., Jin, D.Y., 2007. Human TRBP and PACT directly interact with each other and associate with dicer to facilitate the production of small interfering RNA. J. Biol. Chem. 282, 17649–17657. doi:10.1074/jbc.M611768200

Kota, J., Chivukula, R.R., O'Donnell, K.A., Wentzel, E.A., Montgomery, C.L., Hwang, H.W., Chang, T.C., Vivekanandan, P., Torbenson, M., Clark, K.R., Mendell, J.R., Mendell, J.T., 2009. Therapeutic microRNA Delivery Suppresses Tumorigenesis in a Murine Liver Cancer Model. Cell 137, 1005–1017. doi:10.1016/j.cell.2009.04.021

Krebs, J.E., Goldstein, E.S., Kilpatrick, S.T., 2014. Lewin's Genes XI. Jones & Bartlett Publishers.

Kretz, M., Siprashvili, Z., Chu, C., Webster, D.E., Zehnder, A., Qu, K., Lee, C.S., Flockhart, R.J., Groff, A.F., Chow, J., Johnston, D., Kim, G.E., Spitale, R.C., Flynn, R.A., Zheng, G.X.Y., Aiyer, S., Raj, A., Rinn, J.L., Chang, H.Y., Khavari, P.A., 2013. Control of somatic tissue differentiation by the long non-coding RNA TINCR. Nature 493, 231–5. doi:10.1038/nature11661

Krissinel, E., Henrick, K., 2007. Inference of Macromolecular Assemblies from Crystalline State. J. Mol. Biol. 372, 774–797. doi:10.1016/j.jmb.2007.05.022

Kwak, P.B., Tomari, Y., 2012. The N domain of Argonaute drives duplex unwinding during RISC assembly. Nat. Struct. Mol. Biol. 19, 145–51. doi:10.1038/nsmb.2232

Kwon, S.C., Nguyen, T.A., Choi, Y.G., Jo, M.H., Hohng, S., Kim, V.N., Woo, J.S., 2015. Structure of Human DROSHA. Cell 164, 81–90. doi:10.1016/j.cell.2015.12.019

Kwon, S.C., Yi, H., Eichelbaum, K., Föhr, S., Fischer, B., You, K.T., Castello, A., Krijgsveld, J., Hentze, M.W., Kim, V.N., 2013. The RNA-binding protein repertoire of embryonic stem cells. Nat. Struct. Mol. Biol. 20, 1122–30. doi:10.1038/nsmb.2638

Lander, E.S., Heaford, a, Sheridan, a, Linton, L.M., Birren, B., Subramanian, a, Coulson, a, Nusbaum, C., Zody, M.C., Dunham, a, Baldwin, J., Hunt, a, Devon, K., McMurray, a, Mungall, a, Dewar, K., Doyle, M., Marra, M. a, Fulton, L. a, FitzHugh, W., Chinwalla, a T., Funke, R., Delehaunty, a, Gage, D., Harris, K., Olsen, a, Gibbs, R. a, Fujiyama, a, Howland, J., Toyoda, a, Kann, L., Rosenthal, a, Lehoczky, J., LeVine, R., Rump, a, McEwan, P., Madan, a, McKernan, K., Federspiel, N. a, Abola, a P., Meldrim, J., Mesirov, J.P., Evans, G. a, Roe, B. a, Miranda, C., de La Bastide, M., Morris, W., Bailey, J. a, Naylor, J., Bateman, a, Raymond, C., Rosetti, M., Jones, T. a, Kaspryzk, a, Santos, R., McLysaght, a, Sougnez, C., Smit, a F., Williams, a, Stange-Thomann, N., Stojanovic, N., Felsenfeld, a, Wetterstrand, K. a, Wyman, D., Patrinos, a, Rogers, J., Sulston, J., Ainscough, R., Beck, S.,

Bentley, D., Burton, J., Clee, C., Carter, N., Deadman, R., Deloukas, P., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Jones, M., Lloyd, C., Matthews, L., Mercer, S., Milne, S., Mullikin, J.C., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R.H., Wilson, R.K., Hillier, L.W., McPherson, J.D., Mardis, E.R., Pepin, K.H., Gish, W.R., Chissoe, S.L., Wendl, M.C., Delehaunty, K.D., Miner, T.L., Kramer, J.B., Cook, L.L., Fulton, R.S., Johnson, D.L., Minx, P.J., Clifton, S.W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.F., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Muzny, D.M., Scherer, S.E., Bouck, J.B., Sodergren, E.J., Worley, K.C., Rives, C.M., Gorrell, J.H., Metzker, M.L., Naylor, S.L., Kucherlapati, R.S., Nelson, D.L., Weinstock, G.M., Sakaki, Y., Hattori, M., Yada, T., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D.R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H.M., Dubois, J., Platzer, M., Nyakatura, G., Taudien, S., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Qin, S., Davis, R.W., Proctor, M.J., Myers, R.M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D.R., Olson, M. V, Kaul, R., Shimizu, N., Kawasaki, K., Minoshima, S., Athanasiou, M., Schultz, R., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W.R., Dedhia, N., Blöcker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Batzoglou, S., Birney, E., Bork, P., Brown, D.G., Burge, C.B., Cerutti, L., Chen, H.C., Church, D., Clamp, M., Copley, R.R., Doerks, T., Eddy, S.R., Eichler, E.E., Furey, T.S., Galagan, J., Gilbert, J.G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L.S., Kasif, S., Kennedy, S., Kent, W.J., Kitts, P., Koonin, E. V, Korf, I., Kulp, D., Lancet, D., Lowe, T.M., Mikkelsen, T., Moran, J. V, Mulder, N., Pollara, V.J., Ponting, C.P., Schuler, G., Schultz, J., Slater, G., Stupka, E., Szustakowski, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Wolf, Y.I., Wolfe, K.H., Yang, S.P., Yeh, R.F., Collins, F., Guyer, M.S., Peterson, J., Morgan, M.J., de Jong, P., Catanese, J.J., Osoegawa, K., Shizuya, H., Choi, S., Chen, Y.J., Szustakowki, J., 2001. Initial sequencing and analysis of the human genome. Nature 409, 860–921. doi:10.1038/35057062

Landthaler, M., Yalcin, A., Tuschl, T., 2004. The Human DiGeorge Syndrome Critical Region Gene 8 and Its D. melanogaster Homolog Are Required for miRNA Biogenesis. Curr. Biol. 14, 2162–2167. doi:10.1016/j.cub.2004.11.001

Lanford, R.E., Hildebrandt-Eriksen, E.S., Petri, A., Persson, R., Lindow, M., Munk, M.E., Kauppinen, S., Orum, H., 2010. Therapeutic Silencing of MicroRNA-122 in Primates with Chronic Hepatitis C Virus Infection. Science (80-.). 327, 198–201. doi:10.1126/science.1178178

Lau, C., Bachorik, J.L., Dreyfuss, G., 2009. Gemin5-snRNA interaction reveals an RNA binding

function for WD repeat domains. Nat. Struct. Mol. Biol. 16, 486–91. doi:10.1038/nsmb.1584

Laver, J.D., Li, X., Ray, D., Cook, K.B., Hahn, N.A., Nabeel-Shah, S., Kekis, M., Luo, H., Marsolais, A.J., Fung, K.Y., Hughes, T.R., Westwood, J.T., Sidhu, S.S., Morris, Q., Lipshitz, H.D., Smibert, C.A., 2015. Brain tumor is a sequence-specific RNA-binding protein that directs maternal mRNA clearance during the Drosophila maternal-to-zygotic transition. Genome Biol. 16, 94. doi:10.1186/s13059-015-0659-4

Lazzaretti, D., Tournier, I., Izaurralde, E., 2009. The C-terminal domains of human TNRC6A, TNRC6B, and TNRC6C silence bound transcripts independently of Argonaute proteins. RNA 15, 1059–66. doi:10.1261/rna.1606309

LeCuyer, K.A., Behlen, L.S., Uhlenbeck, O.C., 1996. Mutagenesis of a stacking contact in the MS2 coat protein-RNA complex. EMBO J. 15, 6847–53.

Lee, C.Y., Wilkinson, B.D., Siegrist, S.E., Wharton, R.P., Doe, C.Q., 2006. Brat is a Miranda cargo protein that promotes neuronal differentiation and inhibits neuroblast self-renewal. Dev. Cell 10, 441–449. doi:10.1016/j.devcel.2006.01.017

Lee, R.C., Feinbaum, R.L., Ambros, V., 1993. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75, 843–54. doi:10.1016/0092-8674(93)90529-Y

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Rådmark, O., Kim, S., Kim, V.N., 2003. The nuclear RNase III Drosha initiates microRNA processing. Nature 425, 415–419. doi:10.1038/nature01957

Lee, Y., Jeon, K., Lee, J.T., Kim, S., Kim, V.N., 2002. MicroRNA maturation: Stepwise processing and subcellular localization. EMBO J. 21, 4663–4670. doi:10.1093/emboj/cdf476

Lee, Y.S., Nakahara, K., Pham, J.W., Kim, K., He, Z., Sontheimer, E.J., Carthew, R.W., 2004. Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell 117, 69–81. doi:10.1016/S0092-8674(04)00261-2

Lehmann, R., Nüsslein-Volhard, C., 1987. hunchback, a gene required for segmentation of an anterior and posterior region of the Drosophila embryo. Dev. Biol. 119, 402–17. doi:10.1016/0012-1606(87)90045-5

Leulliot, N., Quevillon-Cheruel, S., Graille, M., van Tilbeurgh, H., Leeper, T.C., Godin, K.S., Edwards, T.E., Sigurdsson, S.T.L., Rozenkrants, N., Nagel, R.J., Ares, M., Varani, G., 2004. A new alpha-helical extension promotes RNA binding by the dsRBD of Rnt1p RNAse III. EMBO J. 23, 2468–77. doi:10.1038/sj.emboj.7600260

Leuschner, P.J.F., Ameres, S.L., Kueng, S., Martinez, J., 2006. Cleavage of the siRNA passenger strand during RISC assembly in human cells. EMBO Rep. 7, 314–320. doi:10.1038/sj.embor.7400637

Lewis, B.P., Shih, I.-H., Jones-Rhoades, M.W., Bartel, D.P., 2003. Prediction of Mammalian MicroRNA Targets that they could have many more regulatory functions than those

uncovered to date (Lagos-Quintana et al. Cell 115, 787–798. doi:10.1016/S0092-8674(03)01018-3

Lewis, H. a, Musunuru, K., Jensen, K.B., Edo, C., Chen, H., Darnell, R.B., Burley, S.K., 2000. Sequence-specific RNA binding by a Nova KH domain: implications for paraneoplastic disease and the fragile X syndrome. Cell 100, 323–32. doi:10.1016/S0092-8674(00)80668-6

Lian, S.L., Li, S., Abadal, G.X., Pauley, B.A., Fritzler, M.J., Chan, E.K.L., 2009. The C-terminal half of human Ago2 binds to multiple GW-rich regions of GW182 and requires GW182 to mediate silencing. RNA 15, 804–13. doi:10.1261/rna.1229409

Liang, C., Wang, Y., Murota, Y., Liu, X., Smith, D., Siomi, M.C., Liu, Q., 2015. TAF11 Assembles the RISC Loading Complex to Enhance RNAi Efficiency. Mol. Cell 59, 807–18. doi:10.1016/j.molcel.2015.07.006

Lim, M.Y.T., Ng, A.W.T., Chou, Y., Lim, T.P., Simcox, A., Tucker-Kellogg, G., Okamura, K., 2016. The Drosophila Dicer-1 Partner Loquacious Enhances miRNA Processing from Hairpins with Unstable Structures at the Dicing Site. Cell Rep. 15, 1–14. doi:10.1016/j.celrep.2016.04.059

Lingel, A., Simon, B., Izaurralde, E., Sattler, M., 2003. Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. Nature 426, 465–469. doi:10.1038/nature02123

Liu, J., Zhang, C., Wang, X.L., Ly, P., Belyi, V., Xu-Monette, Z.Y., Young, K.H., Hu, W., Feng, Z., 2014. E3 ubiquitin ligase TRIM32 negatively regulates tumor suppressor p53 to promote tumorigenesis. Cell Death Differ. 21, 1792–804. doi:10.1038/cdd.2014.121

Liu, Q., Rand, T.A., Kalidas, S., Du, F., Kim, H.-E., Smith, D.P., Wang, X., 2003. R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway. Science 301, 1921–5. doi:10.1126/science.1088710

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25, 402–8. doi:10.1006/meth.2001.1262

Loedige, I., Gaidatzis, D., Sack, R., Meister, G., Filipowicz, W., 2012. The mammalian TRIM-NHL protein TRIM71/LIN-41 is a repressor of mRNA function. Nucleic Acids Res. 41, 518–532. doi:10.1093/nar/gks1032

Loedige, I., Stotz, M., Qamar, S., Kramer, K., Hennig, J., Schubert, T., Löffler, P., Längst, G., Merkl, R., Urlaub, H., Meister, G., 2014. The NHL domain of BRAT is an RNA-binding domain that directly contacts the hunchback mRNA for regulation. Genes Dev. 28, 749–764. doi:10.1101/gad.236513.113

Loop, T., Leemans, R., Stiefel, U., Hermida, L., Egger, B., Xie, F., Primig, M., Certa, U., Fischbach, K.-F., Reichert, H., Hirth, F., 2004. Transcriptional signature of an adult brain tumor in Drosophila. BMC Genomics 5, 24. doi:10.1186/1471-2164-5-24

Lund, E., Güttinger, S., Calado, A., Dahlberg, J.E., Kutay, U., 2004. Nuclear export of microRNA

precursors. Science 303, 95–98. doi:10.1126/science.1090599

Ma, J.-B., Yuan, Y.-R., Meister, G., Pei, Y., Tuschl, T., Patel, D.J., 2005. Structural basis for 5'-end-specific recognition of guide RNA by the A. fulgidus Piwi protein. Nature 434, 666–70. doi:10.1038/nature03514

MacRae, I.J., Ma, E., Zhou, M., Robinson, C. V., Doudna, J.A., 2008. In vitro reconstitution of the human RISC-loading complex. Proc. Natl. Acad. Sci. 105, 512–517. doi:10.1073/pnas.0710869105

Macrae, I.J., Zhou, K., Li, F., Repic, A., Brooks, A.N., Cande, W.Z., Adams, P.D., Doudna, J.A., 2010. Structural Basis for Double-Stranded RNA Processing by Dicer. Science (80-.). 195. doi:10.1126/science.1121638

Maillard, P. V., Ciaudo, C., Marchais, A., Li, Y., Jay, F., Ding, S.W., Voinnet, O., 2013. Antiviral RNA Interference in Mammalian Cells. Science (80-.). 342, 235–238. doi:10.1126/science.1241930

Manche, L., Green, S.R., Schmedt, C., Mathews, M.B., 1992. Interactions between double-stranded RNA regulators and the protein kinase DAI. Mol. Cell. Biol. 12, 5238–48. doi:10.1128/MCB.12.11.5238

Marchetti, G., Reichardt, I., Knoblich, J. a, Besse, F., 2014. The TRIM-NHL protein Brat promotes axon maintenance by repressing src64B expression. J. Neurosci. 34, 13855–64. doi:10.1523/JNEUROSCI.3285-13.2014

Martinez, J., Patkaniowska, A., Urlaub, H., Lührmann, R., Tuschl, T., 2002. Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell 110, 563–74.

Martinez, N.J., Chang, H.-M., Borrajo, J. de R., Gregory, R.I., 2013. The co-chaperones Fkbp4/5 control Argonaute2 expression and facilitate RISC assembly. RNA 19, 1583–93. doi:10.1261/rna.040790.113

Masliah, G., Barraud, P., Allain, F.H.-T., 2013. RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence. Cell. Mol. Life Sci. 70, 1875–95. doi:10.1007/s00018-012-1119-x

Massiah, M.A., Matts, J.A.B., Short, K.M., Simmons, B.N., Singireddy, S., Yi, Z., Cox, T.C., 2007. Solution Structure of the MID1 B-box2 CHC(D/C)C2H2 Zinc-binding Domain: Insights into an Evolutionarily Conserved RING Fold. J. Mol. Biol. 369, 1–10. doi:10.1016/j.jmb.2007.03.017

Mathys, H., Basquin, J., Ozgur, S., Czarnocki-Cieciura, M., Bonneau, F., Aartse, A., Dziembowski, A., Nowotny, M., Conti, E., Filipowicz, W., 2014. Structural and biochemical insights to the role of the CCR4-NOT complex and DDX6 ATPase in microRNA repression. Mol. Cell 54, 751–65. doi:10.1016/j.molcel.2014.03.036

Matranga, C., Tomari, Y., Shin, C., Bartel, D.P., Zamore, P.D., 2005. Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. Cell 123,

607–620. doi:10.1016/j.cell.2005.08.044

Mavrakis, K.J., Wolfe, A.L., Oricchio, E., Palomero, T., de Keersmaecker, K., McJunkin, K., Zuber, J., James, T., Khan, A.A., Leslie, C.S., Parker, J.S., Paddison, P.J., Tam, W., Ferrando, A., Wendel, H.G., 2010. Genome-wide RNA-mediated interference screen identifies miR-19 targets in Notch-induced T-cell acute lymphoblastic leukaemia. Nat Cell Biol 12, 372–379. doi:10.1038/ncb2037

Mazza, C., Segref, A., Mattaj, I.W., Cusack, S., 2002. Large-scale induced fit recognition of an m(7)GpppG cap analogue by the human nuclear cap-binding complex. EMBO J. 21, 5548–57.

McCoy, A.J., 2007. Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr. D. Biol. Crystallogr. 63, 32–41. doi:10.1107/S0907444906045975

Meister, G., 2013. Argonaute proteins: functional insights and emerging roles. Nat. Rev. Genet. 14, 447–59. doi:10.1038/nrg3462

Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., Tuschl, T., 2004. Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. Mol. Cell 15, 185–197. doi:10.1016/j.molcel.2004.07.007

Meister, G., Landthaler, M., Peters, L., Chen, P.Y., Urlaub, H., Lührmann, R., Tuschl, T., 2005. Identification of novel argonaute-associated proteins. Curr. Biol. 15, 2149–55. doi:10.1016/j.cub.2005.10.048

Melo, S., Villanueva, A., Moutinho, C., Davalos, V., Spizzo, R., Ivan, C., Rossi, S., Setien, F., Casanovas, O., Simo-Riudalbas, L., Carmona, J., Carrere, J., Vidal, A., Aytes, A., Puertas, S., Ropero, S., Kalluri, R., Croce, C.M., Calin, G.A., Esteller, M., 2011. Small molecule enoxacin is a cancer-specific growth inhibitor that acts by enhancing TAR RNA-binding protein 2-mediated microRNA processing. Proc. Natl. Acad. Sci. U. S. A. 108, 4394–9. doi:10.1073/pnas.1014720108

Mirkovic-Hösle, M., Förstemann, K., 2014. Transposon defense by endo-siRNAs, piRNAs and somatic pilRNAs in Drosophila: contributions of Loqs-PD and R2D2. PLoS One 9, e84994. doi:10.1371/journal.pone.0084994

Morin, R.D., O'Connor, M.D., Griffith, M., Kuchenbauer, F., Delaney, A., Prabhu, A., Zhao, Y., McDonald, H., Zeng, T., Hirst, M., Eaves, C.J., Marra, M. a, 2008. Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. Genome Res. 18, 610–21. doi:10.1101/gr.7179508

Mourelatos, Z., Dostie, J., Paushkin, S., Sharma, A., Charroux, B., Abel, L., Rappsilber, J., Mann, M., Dreyfuss, G., 2002. miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. Genes Dev. 16, 720–8. doi:10.1101/gad.974702

Nakanishi, K., Weinberg, D.E., Bartel, D.P., Patel, D.J., 2012. Structure of yeast Argonaute with guide RNA. Nature 486, 368–74. doi:10.1038/nature11211

Napoli, C., Lemieux, C., Jorgensen, R., 1990. Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. Plant Cell 2, 279–289. doi:10.1105/tpc.2.4.279

Neumüller, R.A., Betschinger, J., Fischer, A., Bushati, N., Poernbacher, I., Mechtler, K., Cohen, S.M., Knoblich, J.A., 2008. Mei-P26 regulates microRNAs and cell growth in the Drosophila ovarian stem cell lineage. Nature 454, 241–5. doi:10.1038/nature07014

Nishikura, K., 2006. Editor meets silencer: RNA editing and RNA interference. Nat. Rev. Mol. Cell Biol. 7, 919–931. doi:10.1038/nrm2061.Editor

Noland, C.L., Ma, E., Doudna, J.A., 2011. siRNA repositioning for guide strand selection by human Dicer complexes. Mol. Cell 43, 110–21. doi:10.1016/j.molcel.2011.05.028

Nykänen, A., Haley, B., Zamore, P.D., 2001. ATP requirements and small interfering RNA structure in the RNA interference pathway. Cell 107, 309–21. doi:10.1016/S0092-8674(01)00547-5

Obbard, D.J., Jiggins, F.M., Halligan, D.L., Little, T.J., 2006. Natural selection drives extremely rapid evolution in antiviral RNAi genes. Curr. Biol. 16, 580–585. doi:10.1016/j.cub.2006.01.065

Oberstrass, F.C., Lee, A., Stefl, R., Janis, M., Chanfreau, G., Allain, F.H.-T., 2006. Shape-specific recognition in the structure of the Vts1p SAM domain with RNA. Nat. Struct. Mol. Biol. 13, 160–167. doi:10.1038/nsmb1038

Okada, C., Yamashita, E., Lee, S.J., Shibata, S., Katahira, J., Nakagawa, A., Yoneda, Y., Tsukihara, T., 2009. A High-Resolution Structure of the Pre-microRNA Nuclear Export Machinery. Science (80-.). 326, 1275–1279. doi:10.1126/science.1178705

Oubridge, C., Ito, N., Evans, P.R., Teo, C.H., Nagai, K., 1994. Crystal structure at 1.92 A resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. Nature 372, 432–8. doi:10.1038/372432a0

Ozato, K., Shin, D.-M., Chang, T.-H., Morse, H.C., 2008. TRIM family proteins and their emerging roles in innate immunity. Nat. Rev. Immunol. 8, 849–860. doi:10.1038/nri2413

Paddison, P.J., Caudy, A. a, Hannon, G.J., 2002. Stable suppression of gene expression by RNAi in mammalian cells. Proc. Natl. Acad. Sci. U. S. A. 99, 1443–8. doi:10.1073/pnas.032652399

Pak, J., Fire, A., 2007. Distinct populations of primary and secondary effectors during RNAi in C. elegans. Science (80-.). 315, 241–244. doi:10.1126/science.1132839

Pane, A., Wehr, K., Schüpbach, T., 2007. zucchini and squash encode two putative nucleases required for rasiRNA production in the Drosophila germline. Dev. Cell 12, 851–62. doi:10.1016/j.devcel.2007.03.022

Park, J.-E., Heo, I., Tian, Y., Simanshu, D.K., Chang, H., Jee, D., Patel, D.J., Kim, V.N., 2011. Dicer recognizes the 5’ end of RNA for efficient and accurate processing. Nature 475, 201–

205. doi:10.1038/nature10198

Park, J.K., Liu, X., Strauss, T.J., McKearin, D.M., Liu, Q., 2007. The miRNA pathway intrinsically controls self-renewal of Drosophila germline stem cells. Curr. Biol. 17, 533–8. doi:10.1016/j.cub.2007.01.060

Pfaff, J., Hennig, J., Herzog, F., Aebersold, R., Sattler, M., Niessing, D., Meister, G., 2013. Structural features of Argonaute – GW182 protein interactions. Proc. Natl. Acad. Sci. U. S. A. 110, E3770–E3779. doi:10.1073/pnas.1308510110/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1308510110

Pfaff, J., Meister, G., 2013. Argonaute and GW182 proteins: an effective alliance in gene silencing. Biochem. Soc. Trans. 41, 855–860. doi:10.1042/BST20130047

Preus, S., Noer, S.L., Hildebrandt, L.L., Gudnason, D., Birkedal, V., 2015. iSMS: single-molecule FRET microscopy software. Nat. Methods 12, 593–4. doi:10.1038/nmeth.3435

Price, S.R., Evans, P.R., Nagai, K., 1998. Crystal structure of the spliceosomal U2B"-U2A' protein complex bound to a fragment of U2 small nuclear RNA. Nature 394, 645–50. doi:10.1038/29234

Price, S.R., Ito, N., Oubridge, C., Avis, J.M., Nagai, K., 1995. Crystallization of RNA-protein complexes. I. Methods for the large-scale preparation of RNA suitable for crystallographic studies. J. Mol. Biol. 249, 398–408. doi:http://dx.doi.org/10.1006/jmbi.1995.0305

Raheja, R., Liu, Y., Hukkelhoven, E., Yeh, N., Koff, A., 2014. The ability of TRIM3 to induce growth arrest depends on RING-dependent E3 ligase activity. Biochem. J. 458, 537–545. doi:10.1042/BJ20131288

Rand, T.A., Petersen, S., Du, F., Wang, X., 2005. Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. Cell 123, 621–629. doi:10.1016/j.cell.2005.10.020

Rasnik, I., McKinney, S. a, Ha, T., 2006. Nonblinking and long-lasting single-molecule fluorescence imaging. Nat. Methods 3, 891–3. doi:10.1038/nmeth934

Ray, D., Kazan, H., Chan, E.T., Castillo, L.P., Chaudhry, S., Talukder, S., Blencowe, B.J., Morris, Q., Hughes, T.R., 2009. Rapid and systematic analysis of the RNA recognition specificities of RNA-binding proteins. Nat. Biotechnol. 27, 667–670. doi:10.1038/nbt.1550

Reymond, A., Meroni, G., Fantozzi, A., Merla, G., Cairo, S., Luzi, L., Riganelli, D., Zanaria, E., Messali, S., Cainarca, S., Guffanti, A., Minucci, S., Pelicci, P.G., Ballabio, A., 2001. The tripartite motif family identifies cell compartments. EMBO J. 20, 2140–2151. doi:10.1093/emboj/20.9.2140

Rhodes, D.A., De Bono, B., Trowsdale, J., 2005. Relationship between SPRY and B30.2 protein domains. Evolution of a component of immune defence. Immunology 116, 411–417. doi:10.1111/j.1365-2567.2005.02248.x

Rhyu, M.S., Jan, L.Y., Jan, Y.N., 1994. Asymmetric distribution of numb protein during division of the sensory organ precursor cell confers distinct fates to daughter cells. Cell 76, 477–91.

doi:10.1016/0092-8674(94)90112-0

Roeder, R.G., Rutter, W.J., 1970. Specific nucleolar and nucleoplasmic RNA polymerases. Proc. Natl. Acad. Sci. U. S. A. 65, 675–82. doi:10.1073/pnas.65.3.675

Roeder, R.G., Rutter, W.J., 1969. Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. Nature 224, 234–237. doi:10.1038/224234a0

Ruby, T., Bed'Hom, B., Wittzell, H., Morin, V., Oudin, A., Zoorob, R., 2005. Characterisation of a cluster of TRIM-B30.2 genes in the chicken MHC B locus. Immunogenetics 57, 116–128. doi:10.1007/s00251-005-0770-x

Rüdel, S., 2010. Biochemische Studien zur Phosphorylierung humaner Argonaute-Proteine. Ludwig-Maximilians-Universität München.

Rüdel, S., Flatley, A., Weinmann, L., Kremmer, E., Meister, G., 2008. A multifunctional human Argonaute2-specific monoclonal antibody. RNA 14, 1244–1253. doi:10.1261/rna.973808

Rüdel, S., Wang, Y., Lenobel, R., Körner, R., Hsiao, H.H., Urlaub, H., Patel, D., Meister, G., 2011. Phosphorylation of human Argonaute proteins affects small RNA binding. Nucleic Acids Res. 39, 2330–2343. doi:10.1093/nar/gkq1032

Rybak, A., Fuchs, H., Hadian, K., Smirnova, L., Wulczyn, E.A., Michel, G., Nitsch, R., Krappmann, D., Wulczyn, F.G., 2009. The let-7 target gene mouse lin-41 is a stem cell specific E3 ubiquitin ligase for the miRNA pathway protein Ago2. Nat. Cell Biol. 11, 1411–20. doi:10.1038/ncb1987

Ryter, J.M., Schultz, S.C., 1998. Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. EMBO J. 17, 7505–13. doi:10.1093/emboj/17.24.7505

Saito, K., Ishizuka, A., Siomi, H., Siomi, M.C., 2005. Processing of pre-microRNAs by the Dicer-1-Loquacious complex in drosophila cells. PLoS Biol. 3, 1202–1212. doi:10.1371/journal.pbio.0030235

Saito, K., Sakaguchi, Y., Suzuki, T., Suzuki, T., Siomi, H., Siomi, M.C., 2007. Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi- interacting RNAs at their 3' ends. Genes Dev. 21, 1603–8. doi:10.1101/gad.1563607

Sanchez, J.G., Okreglicka, K., Chandrasekaran, V., Welker, J.M., Sundquist, W.I., Pornillos, O., 2014. The tripartite motif coiled-coil is an elongated antiparallel hairpin dimer. Proc. Natl. Acad. Sci. U. S. A. 111, 2494–9. doi:10.1073/pnas.1318962111

Schirle, N.T., MacRae, I.J., 2012. The crystal structure of human Argonaute2. Science 336, 1037–40. doi:10.1126/science.1221551

Schirle, N.T., Sheu-Gruttadauria, J., Chandradoss, S.D., Joo, C., MacRae, I.J., 2015. Water-mediated recognition of t1-adenosine anchors Argonaute2 to microRNA targets. Elife 4, 1–16. doi:10.7554/eLife.07646

Schirle, N.T., Sheu-Gruttadauria, J., MacRae, I.J., 2014. Structural basis for microRNA targeting.

Science 346, 608–13. doi:10.1126/science.1258040

Schraivogel, D., Schindler, S.G., Danner, J., Kremmer, E., Pfaff, J., Hannus, S., Depping, R., Meister, G., 2015. Importin-β facilitates nuclear import of human GW proteins and balances cytoplasmic gene silencing protein levels. Nucleic Acids Res. 43, 7447–61. doi:10.1093/nar/gkv705

Schwamborn, J.C., Berezikov, E., Knoblich, J.A., 2009. The TRIM-NHL protein TRIM32 activates microRNAs and prevents self-renewal in mouse neural progenitors. Cell 136, 913–25. doi:10.1016/j.cell.2008.12.024

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., Selbach, M., 2011. Global quantification of mammalian gene expression control. Nature 473, 337–42. doi:10.1038/nature10098

Schwarz, D.S., Hutvágner, G., Du, T., Xu, Z., Aronin, N., Zamore, P.D., 2003. Asymmetry in the assembly of the RNAi enzyme complex. Cell 115, 199–208. doi:10.1016/S0092-8674(03)00759-1

Schwille, P., Meyer-Almes, F.J., Rigler, R., 1997. Dual-color fluorescence cross-correlation spectroscopy for multicomponent diffusional analysis in solution. Biophys. J. 72, 1878–86. doi:10.1016/S0006-3495(97)78833-7

Seitz, H., Tushir, J.S., Zamore, P.D., 2011. A 5'-uridine amplifies miRNA/miRNA* asymmetry in Drosophila by promoting RNA-induced silencing complex formation. Silence 2, 4. doi:10.1186/1758-907X-2-4

Sengoku, T., Nureki, O., Nakamura, A., Kobayashi, S., Yokoyama, S., 2006. Structural Basis for RNA Unwinding by the DEAD-Box Protein Drosophila Vasa. Cell 125, 287–300. doi:10.1016/j.cell.2006.01.054

Senturia, R., Faller, M., Yin, S., Loo, J.A., Cascio, D., Sawaya, M.R., Hwang, D., Clubb, R.T., Guo, F., 2010. Structure of the dimerization domain of DiGeorge Critical Region 8. Protein Sci. 19, 1354–1365. doi:10.1002/pro.414

Senturia, R., Laganowsky, A., Barr, I., Scheidemantle, B.D., Guo, F., 2012. Dimerization and heme binding are conserved in amphibian and starfish homologues of the microrna processing protein DGCR8. PLoS One 7. doi:10.1371/journal.pone.0039688

Sharma, K., D'Souza, R.C.J., Tyanova, S., Schaab, C., Wi??niewski, J., Cox, J., Mann, M., 2014. Ultradeep Human Phosphoproteome Reveals a Distinct Regulatory Nature of Tyr and Ser/Thr-Based Signaling. Cell Rep. 8, 1583–1594. doi:10.1016/j.celrep.2014.07.036

Short, K.M., Cox, T.C., 2006. Subclassification of the RBCC/TRIM superfamily reveals a novel motif necessary for microtubule binding. J. Biol. Chem. 281, 8970–8980. doi:10.1074/jbc.M512755200

Siomi, H., Matunis, M.J., Michael, W.M., Dreyfuss, G., 1993. The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. Nucleic Acids Res. 21, 1193–8.

doi:10.1093/nar/21.5.1193

Slack, F.J., Ruvkun, G., 1998. A novel repeat domain that is often associated with RING finger and B-box motifs. Trends Biochem. Sci. 23, 474–5. doi:10.1016/S0968-0004(98)01299-7

Smibert, P., Yang, J.-S., Azzam, G., Liu, J.-L., Lai, E.C., 2013. Homeostatic control of Argonaute stability by microRNA availability. Nat. Struct. Mol. Biol. 20, 789–95. doi:10.1038/nsmb.2606

Smith, C.K., Withka, J.M., Regan, L., 1994. A thermodynamic scale for the beta-sheet forming tendencies of the amino acids. Biochemistry 33, 5510–5517. doi:10.1021/bi00184a020

Song, J.-J., Smith, S.K., Hannon, G.J., Joshua-Tor, L., 2004. Crystal structure of Argonaute and its implications for RISC slicer activity. Science 305, 1434–7. doi:10.1126/science.1102514

Sonoda, J., Wharton, R.P., 2001. Drosophila Brain Tumor is a translational repressor. Genes Dev. 15, 762–73. doi:10.1101/gad.870801

Sonoda, J., Wharton, R.P., 1999. Recruitment of nanos to hunchback mRNA by Pumilio. Genes Dev. 13, 2704–2712. doi:10.1101/gad.13.20.2704

Šponer, J., Hobza, P., 2003. Molecular Interactions of Nucleic Acid Bases. A Review of Quantum-Chemical Studies. Collect. Czechoslov. Chem. Commun. 68, 2231–2282. doi:10.1135/cccc20032231

St Johnston, D., Brown, N.H., Gall, J.G., Jantsch, M., 1992. A conserved double-stranded RNA-binding domain. Proc. Natl. Acad. Sci. U. S. A. 89, 10979–83. doi:10.1073/pnas.89.22.10979

Stark, G.R., Kerr, I.M., Williams, B.R., Silverman, R.H., Schreiber, R.D., 1998. How cells respond to interferons. Annu. Rev. Biochem. 67, 227–64. doi:10.1146/annurev.biochem.67.1.227

Stefl, R., Skrisovska, L., Allain, F.H.-T., 2005. RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle. EMBO Rep. 6, 33–38. doi:10.1038/sj.embor.7400325

Stefl, R., Xu, M., Skrisovska, L., Emeson, R.B., Allain, F.H.-T., 2006. Structure and specific RNA binding of ADAR2 double-stranded RNA binding motifs. Structure 14, 345–55. doi:10.1016/j.str.2005.11.013

Stirnimann, C.U., Petsalaki, E., Russell, R.B., Müller, C.W., 2010. WD40 proteins propel cellular networks. Trends Biochem. Sci. 35, 565–74. doi:10.1016/j.tibs.2010.04.003

Subtelny, A.O., Eichhorn, S.W., Chen, G.R., Sive, H., Bartel, D.P., 2014. Poly(A)-tail profiling reveals an embryonic switch in translational control. Nature 508, 66–71. doi:10.1038/nature13007

Szakmary, A., Cox, D.N., Wang, Z., Lin, H., 2005. Regulatory relationship among piwi, pumilio, and bag-of-marbles in Drosophila germline stem cell self-renewal and differentiation. Curr. Biol. 15, 171–178. doi:10.1016/j.cub.2005.01.005

Takimoto, K., Wakiyama, M., Yokoyama, S., 2009. Mammalian GW182 contains multiple Argonaute-binding sites and functions in microRNA-mediated translational repression. RNA 15, 1078–89. doi:10.1261/rna.1363109

Tam, O.H., Aravin, A.A., Stein, P., Girard, A., Murchison, E.P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R.M., Hannon, G.J., 2008. Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature 453, 534–538. doi:nature06904 [pii]\r10.1038/nature06904

Tao, H., Simmons, B.N., Singireddy, S., Jakkidi, M., Short, K.M., Cox, T.C., Massiah, M.A., 2008. Structure of the MID1 tandem B-boxes reveals an interaction reminiscent of intermolecular ring heterodimers. Biochemistry 47, 2450–2457. doi:10.1021/bi7018496

Tautz, D., Lehmann, R., Schnürch, H., Schuh, R., Seifert, E., Kienlin, A., Jones, K., Jäckle, H., 1987. Finger protein of novel structure encoded by hunchback, a second member of the gap class of Drosophila segmentation genes. Nature 327, 383–389. doi:10.1038/327383a0

Tautz, D., Pfeifle, C., 1989. A non-radioactive in situ hybridization method for the localization of specific RNAs in Drosophila embryos reveals translational control of the segmentation gene hunchback. Chromosoma 98, 81–85. doi:10.1007/BF00291041

Tian, Y., Simanshu, D.K., Ma, J.B., Park, J.E., Heo, I., Kim, V.N., Patel, D.J., 2014. A Phosphate-Binding Pocket within the Platform-PAZ-Connector Helix Cassette of Human Dicer. Mol. Cell 53, 606–616. doi:10.1016/j.molcel.2014.01.003

Till, S., Lejeune, E., Thermann, R., Bortfeld, M., Hothorn, M., Enderle, D., Heinrich, C., Hentze, M.W., Ladurner, A.G., 2007. A conserved motif in Argonaute-interacting proteins mediates functional interactions through the Argonaute PIWI domain. Nat. Struct. Mol. Biol. 14, 897–903. doi:10.1038/nsmb1302

Tocchini, C., Keusch, J.J., Miller, S.B., Finger, S., Gut, H., Stadler, M.B., Ciosk, R., 2014. The TRIM-NHL Protein LIN-41 Controls the Onset of Developmental Plasticity in Caenorhabditis elegans. PLoS Genet. 10, e1004533. doi:10.1371/journal.pgen.1004533

Tomari, Y., Du, T., Zamore, P.D., 2007. Sorting of Drosophila small silencing RNAs. Cell 130, 299–308. doi:10.1016/j.cell.2007.05.057

Tomari, Y., Matranga, C., Haley, B., Martinez, N., Zamore, P.D., 2004. A protein sensor for siRNA asymmetry. Science 306, 1377–80. doi:10.1126/science.1102755

Tripsianes, K., Friberg, A., Barrandon, C., Brooks, M., van Tilbeurgh, H., Seraphin, B., Sattler, M., 2014. A novel protein-protein interaction in the RES (REtention and Splicing) complex. J. Biol. Chem. 289, 28640–50. doi:10.1074/jbc.M114.592311

Tritschler, F., Braun, J.E., Eulalio, A., Truffault, V., Izaurralde, E., Weichenrieder, O., 2009. Structural basis for the mutually exclusive anchoring of P body components EDC3 and Tral to the DEAD box protein DDX6/Me31B. Mol. Cell 33, 661–8. doi:10.1016/j.molcel.2009.02.014

Uchil, P.D., Quinlan, B.D., Chan, W.T., Luna, J.M., Mothes, W., 2008. TRIM E3 ligases interfere with early and late stages of the retroviral life cycle. PLoS Pathog. 4. doi:10.1371/journal.ppat.0040016

Vagin, V. V., Sigova, A., Li, C., Seitz, H., Gvozdev, V., Zamore, P.D., 2006. A Distinct Small RNA Pathway Silences Selfish Genetic Elements in the Germline. Science (80-.). 313, 320–324. doi:10.1126/science.1129333

Van Duyne, G.D., Standaert, R.F., Karplus, P.A., Schreiber, S.L., Clardy, J., 1993. Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. J. Mol. Biol. 229, 105–24. doi:10.1006/jmbi.1993.1012

Vilcek, J., 2006. Fifty Years of Interferon Research: Aiming at a Moving Target. Immunity 25, 343–348. doi:10.1016/j.immuni.2006.08.008

Vonrhein, C., Blanc, E., Roversi, P., Bricogne, G., 2007. Automated Structure Solution With autoSHARP, in: Doublié, S. (Ed.), Macromolecular Crystallography Protocols, Volume 2. Humana Press, New Jersey, pp. 215–230. doi:10.1385/1-59745-266-1:215

Walter, T.S., Meier, C., Assenberg, R., Au, K.-F., Ren, J., Verma, A., Nettleship, J.E., Owens, R.J., Stuart, D.I., Grimes, J.M., 2006. Lysine Methylation as a Routine Rescue Strategy for Protein Crystallization. Structure 14, 1617–1622. doi:10.1016/j.str.2006.09.005

Wang, H.-W., Noland, C., Siridechadilok, B., Taylor, D.W., Ma, E., Felderer, K., Doudna, J.A., Nogales, E., 2009. Structural insights into RNA processing by the human RISC-loading complex. Nat. Struct. Mol. Biol. 16, 1148–1153. doi:10.1038/nsmb.1673

Wang, X., McLachlan, J., Zamore, P.D., Hall, T.M.T., 2002. Modular recognition of RNA by a human Pumilio-homology domain. Cell 110, 501–512. doi:10.1016/S0092-8674(02)00873-5

Wang, X., Zamore, P.D., Tanaka Hall, T.M., 2001. Crystal structure of a Pumilio homology domain. Mol. Cell 7, 855–865. doi:10.1016/S1097-2765(01)00229-5

Wang, X.-H., Aliyari, R., Li, W.-X., Li, H.-W., Kim, K., Carthew, R., Atkinson, P., Ding, S.-W., 2006. RNA Interference Directs Innate Immunity Against Viruses in Adult Drosophila. Science (80-.). 312, 452–454.

Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T., Surani, M.A., Sakaki, Y., Sasaki, H., 2008. Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature 453, 539–543. doi:10.1038/nature06908

Wei, W., Ba, Z., Gao, M., Wu, Y., Ma, Y., Amiard, S., White, C.I., Rendtlew Danielsen, J.M., Yang, Y.-G., Qi, Y., 2012. A role for small RNAs in DNA double-strand break repair. Cell 149, 101–12. doi:10.1016/j.cell.2012.03.002

Weinert, C., Grütter, C., Roschitzki-Voser, H., Mittl, P.R.E., Grütter, M.G., 2009. The crystal structure of human pyrin b30.2 domain: implications for mutations associated with familial Mediterranean fever. J. Mol. Biol. 394, 226–36. doi:10.1016/j.jmb.2009.08.059

Weinert, C., Morger, D., Djekic, A., Grütter, M.G., Mittl, P.R.E., 2015. Crystal structure of TRIM20 C-terminal coiled-coil/B30.2 fragment: implications for the recognition of higher order oligomers. Sci. Rep. 5, 10819. doi:10.1038/srep10819

Westover, K.D., Bushnell, D.A., Kornberg, R.D., 2004. Structural Basis of Transcription: Separation of RNA from DNA by RNA Polymerase II. Science (80-.). 303, 1014–1016. doi:10.1126/science.1090839

White, E.K., Moore-Jarrett, T., Ruley, H.E., 2001. PUM2, a novel murine puf protein, and its consensus RNA-binding site. Rna 7, 1855–1866.

Wightman, B., Ha, I., Ruvkun, G., 1993. Posttranscriptional regulation of the heterochronic gene lin- 14 by lin-4 mediates temporal pattern formation in C. elegans. Cell 75, 855–862.

Williams, B.R., 1999. PKR; a sentinel kinase for cellular stress. Oncogene 18, 6112–6120. doi:10.1038/sj.onc.1203127

Wilson, R.C., Tambe, A., Kidwell, M.A., Noland, C.L., Schneider, C.P., Doudna, J.A., 2015. Dicer-TRBP Complex Formation Ensures Accurate Mammalian MicroRNA Biogenesis. Mol. Cell 57, 397–407. doi:10.1016/j.molcel.2014.11.030

Woo, J.S., Suh, H.Y., Park, S.Y., Oh, B.H., 2006. Structural Basis for Protein Recognition by B30.2/SPRY Domains. Mol. Cell 24, 967–976. doi:10.1016/j.molcel.2006.11.009

Woolfson, D.N., Alber, T., 1995. Predicting oligomerization states of coiled coils. Protein Sci. 4, 1596–1607. doi:10.1002/pro.5560040818

Wright, T.R., Bewley, G.C., Sherald, A.F., 1976. The genetics of dopa decarboxylase in Drosophila melanogaster. II. Isolation and characterization of dopa-decarboxylase-deficient mutants and their relationship to the alpha-methyl-dopa-hypersensitive mutants. Genetics 84, 287–310.

Xiao, Q., Komori, H., Lee, C.-Y., 2012. klumpfuss distinguishes stem cells from progenitor cells during asymmetric neuroblast division. Development 139, 2670–2680. doi:10.1242/dev.081687

Yamashita, S., Nagata, T., Kawazoe, M., Takemoto, C., Kigawa, T., Güntert, P., Kobayashi, N., Terada, T., Shirouzu, M., Wakiyama, M., Muto, Y., Yokoyama, S., 2011. Structures of the first and second double-stranded RNA-binding domains of human TAR RNA-binding protein. Protein Sci. 20, 118–130. doi:10.1002/pro.543

Yang, S.W., Chen, H.-Y., Yang, J., Machida, S., Chua, N.-H., Yuan, Y.A., 2010. Structure of Arabidopsis HYPONASTIC LEAVES1 and its molecular implications for miRNA processing. Structure 18, 594–605. doi:10.1016/j.str.2010.02.006

Yang, W., Steitz, T.A., 1995. Recombining the structures of HIV integrase, RuvC and RNase H. Structure 3, 131–4. doi:10.1016/S0969-2126(01)00142-3

Yap, M.W., Nisole, S., Lynch, C., Stoye, J.P., 2004. Trim5alpha protein restricts both HIV-1 and murine leukemia virus. Proc. Natl. Acad. Sci. U. S. A. 101, 10786–91.

doi:10.1073/pnas.0402876101

Ye, X., Huang, N., Liu, Y., Paroo, Z., Huerta, C., Li, P., Chen, S., Liu, Q., Zhang, H., 2011. Structure of C3PO and mechanism of human RISC activation. Nat. Struct. Mol. Biol. 18, 650–7. doi:10.1038/nsmb.2032

Ye, X., Paroo, Z., Liu, Q., 2007. Functional Anatomy of the *Drosophila* MicroRNA-generating Enzyme. J. Biol. Chem. 282, 28373–28378. doi:10.1074/jbc.M705208200

Yi, R., Qin, Y., Macara, I.G., Cullen, B.R., 2003. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev. 17, 3011–6. doi:10.1101/gad.1158803

Yigit, E., Batista, P.J., Bei, Y., Pang, K.M., Chen, C.C.G., Tolia, N.H., Joshua-Tor, L., Mitani, S., Simard, M.J., Mello, C.C., 2006. Analysis of the C. elegans Argonaute Family Reveals that Distinct Argonautes Act Sequentially during RNAi. Cell 127, 747–757. doi:10.1016/j.cell.2006.09.033

Yoda, M., Kawamata, T., Paroo, Z., Ye, X., Iwasaki, S., Liu, Q., Tomari, Y., 2010. ATP-dependent human RISC assembly pathways. Nat. Struct. Mol. Biol. 17, 17–23. doi:10.1038/nsmb.1733

Zhang, H., Kolb, F.A., Brondani, V., Billy, E., Filipowicz, W., 2002. Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. EMBO J. 21, 5875–85. doi:10.1093/emboj/cdf582

Zhang, J., Liu, H., Yao, Q., Yu, X., Chen, Y., Cui, R., Wu, B., Zheng, L., Zuo, J., Huang, Z., Ma, J., Gan, J., 2016. Structural basis for single-stranded RNA recognition and cleavage by C3PO. Nucleic Acids Res. 44, 9494–9504. doi:10.1093/nar/gkw776

List of tables

List of figures

Author contributions

FCCS based Ago loading assay

Fluorescently labeled siRNAs that were ordered for characterization in the FCCS assay were designed according to data from preliminary experiments that were conducted by our collaborator Michael Hannus. Judith Hauptmann performed the cleavage assay of GFP-tagged Ago2 for this thesis. Kerrin Hansen generated stable GFP-Ago2 and GFP-Ago2 Y529E cell lines.

Influence of Ago phosphorylation on siRNA loading

Sabine Rüdel generated the pGTO-b plasmids for GFP-Ago2 Y529 A/E/F and Q that were used for the generation of stable cell lines.

***In vitro* characterization of the Loqs-dmDcr-1 interaction interface**

FCCS titration measurements were performed by Kerrin Hansen with supplied material at the Intana bioscience GmbH.

DmDcr-1 interaction in context of Loqs-dsRBD3 dimerization

Kevin Kramm, Alexander Gust and Dina Grohmann performed and optimized single molecule bleaching experiments (SiMPull) with supplied material and designed Figure 2.24.

Purification of the Brat-NHL-RNA complex

Mathias Stotz and Inga Lödige provided purified Brat-NHL protein and consensus RNA for crystallography experiments.

Acknowledgements

First, I want to thank Gunter Meister for giving me the opportunity to work on my PhD Thesis in his group. I want to thank you for your constant support and your advice within my research projects and funding of my experiments. From your mentorship I could benefit a lot. Thank you very much Gunter!

Additionally I want to thank Michael and Stefan Hannus for funding of my position within Gunter's group. Thank you Michael for the supervision of my Ago-projects and Stefan for leading me into the exciting field of fluorescence spectroscopy.

I am deeply grateful to Nora and Thomas Treiber who constantly supported me within all of my crystallographic projects during working hours and even beyond at the weekends. I wish you all the best for your family and your further scientific careers.

Many thanks to all the nice people who shared the lab with me. First I want to thank Daniele Hasler who became more than just a colleague or a good friend for me as I could win you for my witness to marriage. Thank you Daniele for all the helpful discussions and the constant support related or unrelated from my work.

In a similar way I want to thank Ludwig Wankerl for his support concerning my paper and for all the nice time we had inside and outside the lab. I will also miss our little "arguments".

A great part for my crystallographic knowledge and progress of the projects I owe to Mathias Stotz. I am very happy that you started your postdoc in Gunter's lab and you found the time for all the useful scientific discussions.

All the others from the lab I want to thank for the great atmosphere and for the nice time which we could spend together. Thanks of course also to all my practical students, who supported my work.

I also want to thank my collaborators from the AG Grohmann for their great work for my publication, which would not have been so easy to publish without. Thank you Dina, Alex and Kevin for your great support!

Finally I want to thank people from outside the lab. A really great "thank you" goes to my lovely wife Nadja and my beautiful little daughter Sophia for their great support at home and for taking my mind off work. Thank you for always being there when I need you! And also many thanks to my parents and my brother for your constant support, which was crucial for me for coming this far.

Thank you all very much!